

GENE REGULATION IN TRANSGENIC ANIMALS USING A TRANSPOSON-BASED VECTOR

10 The U.S. Government has certain rights in this invention. The development of this invention was partially funded by the United States Government under a HATCH grant from the United States Department of Agriculture, partially funded by the United States Government with Formula 1433 funds from the United States Department of Agriculture and partially funded by the United States Government
15 under contract DAAD 19-02016 awarded by the Army.

FIELD OF THE INVENTION

20 The present invention relates generally to cell-specific gene regulation in transgenic animals. Animals may be made transgenic through administration of a transposon-based vector through any method of administration including pronuclear injection, or intraembryonic, intratesticular, intraoviductal or intravenous administration. These transgenic animals contain the gene of interest in all cells, including germ cells. Animals may also be made transgenic by targeting specific cells for uptake and gene incorporation of the transposon-based vectors. Stable
25 incorporation of a gene of interest into cells of the transgenic animals is demonstrated by expression of the gene of interest in a cell, wherein expression is regulated by a promoter sequence. The promoter sequence may be provided as a transgene along with the gene of interest or may be endogenous to the cell. The promoter sequence may be constitutive or inducible, wherein inducible promoters include tissue-specific
30 promoters, developmentally regulated promoters and chemically inducible promoters.

BACKGROUND OF THE INVENTION

35 Transgenic animals are desirable for a variety of reasons, including their potential as biological factories to produce desired molecules for pharmaceutical, diagnostic and industrial uses. This potential is attractive to the industry due to the inadequate capacity in facilities used for recombinant production of desired molecules and the increasing demand by the pharmaceutical industry for use of these facilities. Numerous attempts to produce transgenic animals have met several problems,

including low rates of gene incorporation and unstable gene incorporation. Accordingly, improved gene technologies are needed for the development of transgenic animals for the production of desired molecules.

Improved gene delivery technologies are also needed for the treatment of disease in animals and humans. Many diseases and conditions can be treated with gene-delivery technologies, which provide a gene of interest to a patient suffering from the disease or the condition. An example of such disease is Type 1 diabetes. Type 1 diabetes is an autoimmune disease that ultimately results in destruction of the insulin producing β -cells in the pancreas. Although patients with Type 1 diabetes may be treated adequately with insulin injections or insulin pumps, these therapies are only partially effective. Insulin replacement, such as via insulin injection or pump administration, cannot fully reverse the defect in the vascular endothelium found in the hyperglycemic state (Pieper et al., 1996. *Diabetes Res. Clin. Pract. Suppl.* S157-S162). In addition, hyper- and hypoglycemia occurs frequently despite intensive home blood glucose monitoring. Finally, careful dietary constraints are needed to maintain an adequate ratio of consumed calories consumed. This often causes major psychosocial stress for many diabetic patients. Development of gene therapies providing delivery of the insulin gene into the pancreas of diabetic patients could overcome many of these problems and result in improved life expectancy and quality of life.

Several of the prior art gene delivery technologies employed viruses that are associated with potentially undesirable side effects and safety concerns. The majority of current gene-delivery technologies useful for gene therapy rely on virus-based delivery vectors, such as adeno and adeno-associated viruses, retroviruses, and other viruses, which have been attenuated to no longer replicate. (Kay, M.A., et al. 2001. *Nature Medicine* 7:33-40).

There are multiple problems associated with the use of viral vectors. First, they are not tissue-specific. In fact, a gene therapy trial using adenovirus was recently halted because the vector was present in the patient's sperm (Gene trial to proceed despite fears that therapy could change child's genetic makeup. *The New York Times*, December 23, 2001). Second, viral vectors are likely to be transiently incorporated, which necessitates re-treating a patient at specified time intervals. (Kay, M.A., et al. 2001. *Nature Medicine* 7:33-40). Third, there is a concern that a viral-based vector could revert to its virulent form and cause disease. Fourth, viral-based vectors require a dividing cell for stable integration. Fifth, viral-based vectors indiscriminately integrate into various cells and tissues, which can result in undesirable germline integration. Sixth, the required high titers needed to achieve the desired effect have resulted in the death of one patient and they are believed to be

responsible for induction of cancer in a separate study. (Science, News of the Week, October 4, 2002).

Accordingly, what is needed is a new vector to produce transgenic animals and humans with stably incorporated genes, which vector does not cause disease or other unwanted side effects. There is also a need for DNA constructs that would be stably incorporated into the tissues and cells of animals and humans, including cells in the resting state, which are not replicating. There is a further recognized need in the art for DNA constructs capable of delivering genes to specific tissues and cells of animals and humans.

When incorporating a gene of interest into an animal for the production of a desired protein or when incorporating a gene of interest in an animal or human for the treatment of a disease, it is often desirable to selectively activate incorporated genes using inducible promoters. These inducible promoters are regulated by substances either produced or recognized by the transcription control elements within the cell in which the gene is incorporated. In many instances, control of gene expression is desired in transgenic animals or humans so that incorporated genes are selectively activated at desired times and/or under the influence of specific substances. Accordingly, what is needed is a means to selectively activate genes introduced into the genome of cells of a transgenic animal or human. This can be taken a step further to cause incorporation to be tissue-specific, which prevents widespread gene incorporation throughout a patient's body (animal or human). This decreases the amount of DNA needed for a treatment, decreases the chance of incorporation in gametes, and targets gene delivery, incorporation, and expression to the desired tissue where the gene is needed to function.

25

SUMMARY OF THE INVENTION

The present invention addresses the problems described above by providing new, effective and efficient compositions for producing transgenic animals and for treating disease in animals or humans. Transgenic animals include all egg-laying animals and milk-producing animals. Transgenic animals further include but are not limited to avians, fish, amphibians, reptiles, insects, mammals and humans. In a preferred embodiment, the animal is an avian animal. In another preferred embodiment, the animal is a milk-producing animal, including but not limited to bovine, porcine, ovine and equine animals. Animals are made transgenic through administration of a composition comprising a transposon-based vector designed for stable incorporation of a gene of interest for production of a desired protein, together

with an acceptable carrier. A transfection reagent is optionally added to the composition before administration.

The transposon-based vectors of the present invention include a transposase, operably-linked to a first promoter, and a coding sequence for a protein or peptide of interest operably-linked to a second promoter, wherein the coding sequence for the protein or peptide of interest and its operably-linked promoter are flanked by transposase insertion sequences recognized by the transposase. The transposon-based vector also includes the following characteristics: a) one or more modified Kozak sequences comprising ACCATG (SEQ ID NO:13) at the 3' end of the first promoter to enhance expression of the transposase; b) modifications of the codons for the first several N-terminal amino acids of the transposase, wherein the nucleotide at the third base position of each codon was changed to an A or a T without changing the corresponding amino acid; c) addition of one or more stop codons to enhance the termination of transposase synthesis; and/or, d) addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene.

Use of the compositions of the present invention results in highly efficient and stable incorporation of a gene of interest into the genome of transfected animals. For example, transgenic avians have been mated and produce transgenic progeny in the G1 generation. The transgenic progeny have been mated and produce transgenic progeny in the G2 generation.

The present invention also provides for tissue-specific incorporation and/or expression of a gene of interest. Tissue-specific incorporation of a gene of interest may be achieved by placing the transposase gene under the control of a tissue-specific promoter, whereas tissue-specific expression of a gene of interest may be achieved by placing the gene of interest under the control of a tissue-specific promoter. In some embodiments, the gene of interest is transcribed under the influence of an ovalbumin, or other oviduct specific, promoter. Linking the gene of interest to an oviduct specific promoter in an egg-laying animal results in synthesis of a desired molecule and deposition of the desired molecule in a developing egg. The present invention further provides for stable incorporation and expression of genes in the epithelial cells of the mammary gland in milk-producing animals. Transcription of the gene of interest in the epithelial cells of the mammary gland results in synthesis of a desired molecule

and deposition of the desired molecule in the milk. A preferred molecule is a protein.

In some embodiments, the desired molecule deposited in the milk is an antiviral protein, an antibody, or a serum protein.

In other embodiments, specific incorporation of the proinsulin gene into liver cells of a diabetic animal results in the improvement of the animal's condition. Such improvement is achieved by placing a transposase gene under the control of a liver-specific promoter, which drives integration of the gene of interest in liver cells of the diabetic animal.

The present invention advantageously produces a high number of transgenic animals having a gene of interest stably incorporated. These transgenic animals successfully pass the desired gene to their progeny. The transgenic animals of the present invention also produce large amounts of a desired molecule encoded by the transgene. Transgenic egg-laying animals, particularly avians, produce large amounts of a desired protein that is deposited in the egg for rapid harvest and purification.

Transgenic milk-producing animals produce large amounts of a desired protein that is deposited in the milk for rapid harvest and purification.

Any desired gene may be incorporated into the novel transposon-based vectors of the present invention in order to synthesize a desired molecule in the transgenic animals. Proteins, peptides and nucleic acids are preferred desired molecules to be produced by the transgenic animals of the present invention. Particularly preferred proteins are antibody proteins.

This invention provides a composition useful for the production of transgenic hens capable of producing substantially high amounts of a desired protein or peptide. Entire flocks of transgenic birds may be developed very quickly in order to produce industrial amounts of desired molecules. The present invention solves the problems inherent in the inadequate capacity of fermentation facilities used for bacterial production of molecules and provides a more efficient and economical way to produce desired molecules. Accordingly, the present invention provides a means to produce large amounts of therapeutic, diagnostic and reagent molecules.

Transgenic chickens are excellent in terms of convenience and efficiency of manufacturing molecules such as proteins and peptides. Starting with a single transgenic rooster, thousands of transgenic offspring can be produced within a year. (In principle, up to forty million offspring could be produced in just three

generations). Each transgenic female is expected to lay at least 250 eggs/year, each potentially containing hundreds of milligrams of the selected protein. Flocks of chickens numbering in the hundreds of thousands are readily handled through established commercial systems. The technologies for obtaining eggs and 5 fractionating them are also well known and widely accepted. Thus, for each therapeutic, diagnostic, or other protein of interest, large amounts of a substantially pure material can be produced at relatively low incremental cost.

A wide range of recombinant peptides and proteins can be produced in transgenic egg-laying animals and milk-producing animals. Enzymes, hormones, 10 antibodies, growth factors, serum proteins, commodity proteins, biological response modifiers, peptides and designed proteins may all be made through practice of the present invention. For example, rough estimates suggest that it is possible to produce in bulk growth hormone, insulin, or Factor VIII, and deposit them in transgenic egg whites, for an incremental cost in the order of one dollar per gram. At such prices it 15 is feasible to consider administering such medical agents by inhalation or even orally, instead of through injection. Even if bioavailability rates through these avenues were low, the cost of a much higher effective-dose would not be prohibitive.

In one embodiment, the egg-laying transgenic animal is an avian. The method of the present invention may be used in avians including Ratites, Psittaciformes, 20 Falconiformes, Piciformes, Strigiformes, Passeriformes, Coraciformes, Ralliformes, Cuculiformes, Columbiformes, Galliformes, Anseriformes, and Herodiones. Preferably, the egg-laying transgenic animal is a poultry bird. More preferably, the bird is a chicken, turkey, duck, goose or quail. Another preferred bird is a ratite, such as, an emu, an ostrich, a rhea, or a cassowary. Other preferred birds are partridge, 25 pheasant, kiwi, parrot, parakeet, macaw, falcon, eagle, hawk, pigeon, cockatoo, song birds, jay bird, blackbird, finch, warbler, canary, toucan, mynah, or sparrow.

In another embodiment, the transgenic animal is a milk-producing animal, including but not limited to bovine, ovine, porcine, equine, and primate animals. Milk-producing animals include but are not limited to cows, goats, horses, pigs, 30 buffalo, rabbits, non-human primates, and humans.

Accordingly, it is an object of the present invention to provide novel transposon-based vectors.

It is another object of the present invention to provide novel transposon-based vectors that encode for the production of desired proteins or peptides in cells.

It is an object of the present invention to produce transgenic animals through administration of a transposon-based vector.

5 Another object of the present invention is to produce transgenic animals through administration of a transposon-based vector, wherein the transgenic animals produce desired proteins or peptides.

10 Yet another object of the present invention is to produce transgenic animals through administration of a transposon-based vector, wherein the transgenic animals produce desired proteins or peptides and deposit the proteins or peptides in eggs or milk.

15 It is a further object of the present invention to produce transgenic animals through intraembryonic, intratesticular or intraoviductal administration of a transposon-based vector.

20 It is further an object of the present invention to provide a method to produce transgenic animals through administration of a transposon-based vector that are capable of producing transgenic progeny.

25 Yet another object of the present invention is to provide a method to produce transgenic animals through administration of a transposon-based vector that are capable of producing a desired molecule, such as a protein, peptide or nucleic acid.

Another object of the present invention is to provide a method to produce transgenic animals through administration of a transposon-based vector, wherein such administration results in modulation of endogenous gene expression.

30 It is another object of the present invention to provide transposon-vectors useful for cell- or tissue-specific expression of a gene of interest in an animal or human with the purpose of gene therapy.

It is yet another object of the present invention to provide a method to produce transgenic avians through administration of a transposon-based vector that are capable of producing proteins, peptides or nucleic acids.

35 It is another object of the present invention to produce transgenic animals through administration of a transposon-based vector encoding an antibody or a fragment thereof.

Still another object of the present invention is to provide a method to produce transgenic avians through administration of a transposon-based vector that are capable of producing proteins or peptides and depositing these proteins or peptides in the egg.

Another object of the present invention is to provide transgenic avians that 5 contain a stably incorporated transgene.

Still another object of the present invention is to provide eggs containing desired proteins or peptides encoded by a transgene incorporated into the transgenic avian that produces the egg.

It is further an object of the present invention to provide a method to produce 10 transgenic milk-producing animals through administration of a transposon-based vector that are capable of producing proteins, peptides or nucleic acids.

Still another object of the present invention is to provide a method to produce transgenic milk-producing animals through administration of a transposon-based vector that are capable of producing proteins or peptides and depositing these proteins 15 or peptides in their milk.

Another object of the present invention is to provide transgenic milk-producing animals that contain a stably incorporated transgene.

Another object of the present invention is to provide transgenic milk-producing animals that are capable of producing proteins or peptides and depositing 20 these proteins or peptides in their milk.

Yet another object of the present invention is to provide milk containing desired molecules encoded by a transgene incorporated into the transgenic milk-producing animals that produce the milk.

Still another object of the present invention is to provide milk containing 25 desired proteins or peptides encoded by a transgene incorporated into the transgenic milk-producing animals that produce the milk.

A further object of the present invention to provide a method to produce transgenic sperm through administration of a transposon-based vector to an animal.

A further object of the present invention to provide transgenic sperm that 30 contain a stably incorporated transgene.

An advantage of the present invention is that transgenic animals are produced with higher efficiencies than observed in the prior art.

Another advantage of the present invention is that these transgenic animals possess high copy numbers of the transgene.

Another advantage of the present invention is that the transgenic animals produce large amounts of desired molecules encoded by the transgene.

5 Still another advantage of the present invention is that desired molecules are produced by the transgenic animals much more efficiently and economically than prior art methods, thereby providing a means for large scale production of desired molecules, particularly proteins and peptides.

10 These and other objects, features and advantages of the present invention will become apparent after a review of the following detailed description of the disclosed embodiments and claims.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts schematically a transposon-based vector containing a
15 transposase operably linked to a first promoter and a gene of interest operably-linked
to a second promoter, wherein the gene of interest and its operably-linked promoter
are flanked by insertion sequences (IS) recognized by the transposase. "Pro"
designates a promoter. In this and subsequent figures, the size of the actual nucleotide
sequence is not necessarily proportionate to the box representing that sequence.

20

Figure 2 depicts schematically a transposon-based vector for targeting
deposition of a polypeptide in an egg white wherein Ov pro is the ovalbumin
25 promoter, Ov protein is the ovalbumin protein and PolyA is a polyadenylation
sequence. The TAG sequence includes a spacer, the gp41 hairpin loop from HIV I
and a protein cleavage site.

Figure 3 depicts schematically a transposon-based vector for targeting
deposition of a polypeptide in an egg white wherein Ovo pro is the ovomucoid
30 promoter and Ovo SS is the ovomucoid signal sequence. The TAG sequence includes
a spacer, the gp41 hairpin loop from HIV I and a protein cleavage site.

Figure 4 depicts schematically a transposon-based vector for targeting deposition of a polypeptide in an egg yolk wherein Vit pro is the vitellogenin promoter and Vit targ is the vitellogenin targeting sequence.

5 Figure 5 depicts schematically a transposon-based vector for expression of antibody heavy and light chains. Prepro indicates a prepro sequence from cecropin and pro indicates a pro sequence from cecropin.

10 Figure 6 depicts schematically a transposon-based vector for expression of antibody heavy and light chains. Ent indicates an enterokinase cleavage sequence.

15 Figure 7 depicts schematically egg white targeted expression of antibody heavy and light chains from one vector in either tail-to-tail (Figure 7A) or tail-to-head (Figure 7B) configuration. In the tail-to-tail configuration, the ovalbumin signal sequence adjacent to the gene for the light chain contains on its 3' end an enterokinase cleavage site (not shown) to allow cleavage of the signal sequence from the light chain, and the ovalbumin signal sequence adjacent to the gene for the heavy chain contains on its 5' end an enterokinase cleavage site (not shown) to allow cleavage of the signal sequence from the heavy chain. In the tail-to-head configuration, the 20 ovalbumin signal sequence adjacent to the gene for the heavy chain and the light chain contains on its 3' end an enterokinase cleavage site (not shown) to allow cleavage of the signal sequence from the heavy or light chain.

DETAILED DESCRIPTION OF THE INVENTION

25 The present invention provides a new, effective and efficient method of producing transgenic animals, particularly egg-laying animals and milk-producing animals, through administration of a composition comprising a transposon-based vector designed for stable incorporation of a gene of interest for production of a desired molecule.

30 **Definitions**

It is to be understood that as used in the specification and in the claims, "a" or "an" can mean one or more, depending upon the context in which it is used. Thus, for example, reference to "a cell" can mean that at least one cell can be utilized.

The term "antibody" is used interchangeably with the term "immunoglobulin" and is defined herein as a protein synthesized by an animal or a cell of the immune system in response to the presence of a foreign substance commonly referred to as an "antigen" or an "immunogen". The term antibody includes fragments of antibodies.

5 Antibodies are characterized by specific affinity to a site on the antigen, wherein the site is referred to an "antigenic determinant" or an "epitope". Antigens can be naturally occurring or artificially engineered. Artificially engineered antigens include but are not limited to small molecules, such as small peptides, attached to haptens such as macromolecules, for example proteins, nucleic acids, or polysaccharides.

10 Artificially designed or engineered variants of naturally occurring antibodies and artificially designed or engineered antibodies not occurring in nature are all included in the current definition. Such variants include conservatively substituted amino acids and other forms of substitution as described in the section concerning proteins and polypeptides.

15 As used herein, the term "egg-laying animal" includes all amniotes such as birds, turtles, lizards and monotremes. Monotremes are egg-laying mammals and include the platypus and echidna. The term "bird" or "fowl," as used herein, is defined as a member of the Aves class of animals which are characterized as warm-blooded, egg-laying vertebrates primarily adapted for flying. Avians include, without limitation, Ratites, Psittaciformes, Falconiformes, Piciformes, Strigiformes, Passeriformes, Coraciformes, Ralliformes, Cuculiformes, Columbiformes, Galliformes, Anseriformes, and Herodiones. The term "Ratite," as used herein, is defined as a group of flightless, mostly large, running birds comprising several orders and including the emus, ostriches, kiwis, and cassowaries. The term "Psittaciformes", as used herein, includes parrots and refers to a monofamilial order of birds that exhibit zygodactylism and have a strong hooked bill. A "parrot" is defined as any member of the avian family Psittacidae (the single family of the Psittaciformes), distinguished by the short, stout, strongly hooked beak. The term "chicken" as used herein denotes chickens used for table egg production, such as egg-type chickens, chickens reared for public meat consumption, or broilers, and chickens reared for both egg and meat production ("dual-purpose" chickens). The term "chicken" also denotes chickens produced by primary breeder companies, or chickens that are the parents,

grandparents, great-grandparents, etc. of those chickens reared for public table egg, meat, or table egg and meat consumption.

The term "egg" is defined herein as a large female sex cell enclosed in a porous, calcareous or leathery shell, produced by birds and reptiles. The term "ovum"

5 is defined as a female gamete, and is also known as an egg. Therefore, egg production in all animals other than birds and reptiles, as used herein, is defined as the production and discharge of an ovum from an ovary, or "ovulation". Accordingly, it is to be understood that the term "egg" as used herein is defined as a large female sex cell enclosed in a porous, calcareous or leathery shell, when a bird or reptile produces

10 it, or it is an ovum when it is produced by all other animals.

The term "milk-producing animal" refers herein to mammals including, but not limited to, bovine, ovine, porcine, equine, and primate animals. Milk-producing animals include but are not limited to cows, llamas, camels, goats, reindeer, zebu, water buffalo, yak, horses, pigs, rabbits, non-human primates, and humans.

15 The term "gene" is defined herein to include a coding region for a protein, peptide or polypeptide.

The term "vector" is used interchangeably with the terms "construct", "DNA construct" and "genetic construct" to denote synthetic nucleotide sequences used for manipulation of genetic material, including but not limited to cloning, subcloning,

20 sequencing, or introduction of exogenous genetic material into cells, tissues or organisms, such as birds. It is understood by one skilled in the art that vectors may contain synthetic DNA sequences, naturally occurring DNA sequences, or both. The vectors of the present invention are transposon-based vectors as described herein.

When referring to two nucleotide sequences, one being a regulatory sequence, 25 the term "operably-linked" is defined herein to mean that the two sequences are associated in a manner that allows the regulatory sequence to affect expression of the other nucleotide sequence. It is not required that the operably-linked sequences be directly adjacent to one another with no intervening sequence(s).

The term "regulatory sequence" is defined herein as including promoters, 30 enhancers and other expression control elements such as polyadenylation sequences, matrix attachment sites, insulator regions for expression of multiple genes on a single construct, ribosome entry/attachment sites, introns that are able to enhance expression, and silencers.

Transposon-Based Vectors

While not wanting to be bound by the following statement, it is believed that the nature of the DNA construct is an important factor in successfully producing transgenic animals. The "standard" types of plasmid and viral vectors that have 5 previously been almost universally used for transgenic work in all species, especially avians, have low efficiencies and may constitute a major reason for the low rates of transformation previously observed. The DNA (or RNA) constructs previously used often do not integrate into the host DNA, or integrate only at low frequencies. Other factors may have also played a part, such as poor entry of the vector into target cells.

10 The present invention provides transposon-based vectors that can be administered to an animal that overcome the prior art problems relating to low transgene integration frequencies. Two preferred transposon-based vectors of the present invention in which a transposase, gene of interest and other polynucleotide sequences may be introduced are termed pTnMCS (SEQ ID NO:36) and pTnMod (SEQ ID NO:1).

15 The transposon-based vectors of the present invention produce integration frequencies an order of magnitude greater than has been achieved with previous vectors. More specifically, intratesticular injections performed with a prior art transposon-based vector (described in U.S. Patent No. 5,719,055) resulted in 41% sperm positive roosters whereas intratesticular injections performed with the novel 20 transposon-based vectors of the present invention resulted in 77% sperm positive roosters. Actual frequencies of integration were estimated by either or both comparative strength of the PCR signal from the sperm and histological evaluation of the testes and sperm by quantitative PCR.

The transposon-based vectors of the present invention include a transposase 25 gene operably-linked to a first promoter, and a coding sequence for a desired protein or peptide operably-linked to a second promoter, wherein the coding sequence for the desired protein or peptide and its operably-linked promoter are flanked by transposase insertion sequences recognized by the transposase. The transposon-based vector also includes one or more of the following characteristics: a) one or more modified Kozak 30 sequences comprising ACCATG (SEQ ID NO:13) at the 3' end of the first promoter to enhance expression of the transposase; b) modifications of the codons for the first several N-terminal amino acids of the transposase, wherein the third base of each codon was changed to an A or a T without changing the corresponding amino acid; c)

addition of one or more stop codons to enhance the termination of transposase synthesis; and, d) addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene. Figure 1 shows a schematic representation of several components of the transposon-based vector. The 5 present invention further includes vectors containing more than one gene of interest, wherein a second or subsequent gene of interest is operably-linked to the second promoter or to a different promoter. It is also to be understood that the transposon-based vectors shown in the Figures are representational of the present invention and that the order of the vector elements may be different than that shown in the Figures, 10 that the elements may be present in various orientations, and that the vectors may contain additional elements not shown in the Figures.

Transposases and Insertion Sequences

In a further embodiment of the present invention, the transposase found in the transposase-based vector is an altered target site (ATS) transposase and the insertion 15 sequences are those recognized by the ATS transposase. However, the transposase located in the transposase-based vectors is not limited to a modified ATS transposase and can be derived from any transposase. Transposases known in the prior art include those found in AC7, Tn5SEQ1, Tn916, Tn951, Tn1721, Tn 2410, Tn1681, Tn1, Tn2, Tn3, Tn4, Tn5, Tn6, Tn9, Tn10, Tn30, Tn101, Tn903, Tn501, Tn1000 ($\gamma\delta$), Tn1681, 20 Tn2901, AC transposons, Mp transposons, Spm transposons, En transposons, Dotted transposons, Mu transposons, Ds transposons, dSpm transposons and I transposons. According to the present invention, these transposases and their regulatory sequences are modified for improved functioning as follows: a) the addition one or more modified Kozak sequences comprising ACCATG (SEQ ID NO:13) at the 3' end of 25 the promoter operably-linked to the transposase; b) a change of the codons for the first several amino acids of the transposase, wherein the third base of each codon was changed to an A or a T without changing the corresponding amino acid; c) the addition of one or more stop codons to enhance the termination of transposase synthesis; and/or, d) the addition of an effective polyA sequence operably-linked to 30 the transposase to further enhance expression of the transposase gene.

Although not wanting to be bound by the following statement, it is believed that the modifications of the first several N-terminal codons of the transposase gene increase transcription of the transposase gene, in part, by increasing strand

dissociation. It is preferable that between approximately 1 and 20, more preferably 3 and 15, and most preferably between 4 and 12 of the first N-terminal codons of the transposase are modified such that the third base of each codon is changed to an A or a T without changing the encoded amino acid. In one embodiment, the first ten N-5 terminal codons of the transposase gene are modified in this manner. It is also preferred that the transposase contain mutations that make it less specific for preferred insertion sites and thus increases the rate of transgene insertion as discussed in U.S. Patent No. 5,719,055.

In some embodiments, the transposon-based vectors are optimized for 10 expression in a particular host by changing the methylation patterns of the vector DNA. For example, prokaryotic methylation may be reduced by using a methylation deficient organism for production of the transposon-based vector. The transposon-based vectors may also be methylated to resemble eukaryotic DNA for expression in a eukaryotic host.

15 Transposases and insertion sequences from other analogous eukaryotic transposon-based vectors that can also be modified and used are, for example, the *Drosophila* P element derived vectors disclosed in U.S. Patent No. 6,291,243; the *Drosophila* mariner element described in Sherman et al. (1998); or the sleeping beauty transposon. See also Hackett et al. (1999); D. Lampe et al., 1999. Proc. Natl. Acad. 20 Sci. USA, 96:11428-11433; S. Fischer et al., 2001. Proc. Natl. Acad. Sci. USA, 98:6759-6764; L. Zagoraiou et al., 2001. Proc. Natl. Acad. Sci. USA, 98:11474-11478; and D. Berg et al. (Eds.), *Mobile DNA*, Amer. Soc. Microbiol. (Washington, D.C., 1989). However, it should be noted that bacterial transposon-based elements are preferred, as there is less likelihood that a eukaryotic transposase in the recipient 25 species will recognize prokaryotic insertion sequences bracketing the transgene.

Many transposases recognize different insertion sequences, and therefore, it is to be understood that a transposase-based vector will contain insertion sequences recognized by the particular transposase also found in the transposase-based vector. In a preferred embodiment of the invention, the insertion sequences have been 30 shortened to about 70 base pairs in length as compared to those found in wild-type transposons that typically contain insertion sequences of well over 100 base pairs.

While the examples provided below incorporate a "cut and insert" *Tn10* based vector that is destroyed following the insertion event, the present invention also

encompasses the use of a "rolling replication" type transposon-based vector. Use of a rolling replication type transposon allows multiple copies of the transposon/transgene to be made from a single transgene construct and the copies inserted. This type of transposon-based system thereby provides for insertion of multiple copies of a transgene into a single genome. A rolling replication type transposon-based vector may be preferred when the promoter operably-linked to gene of interest is endogenous to the host cell and present in a high copy number or highly expressed. However, use of a rolling replication system may require tight control to limit the insertion events to non-lethal levels. Tn1, Tn2, Tn3, Tn4, Tn5, Tn9, Tn21, Tn501, Tn551, Tn951, Tn1721, Tn2410 and Tn2603 are examples of a rolling replication type transposon, although Tn5 could be both a rolling replication and a cut and insert type transposon.

Stop Codons and PolyA Sequences

In one embodiment, the transposon-based vector contains two stop codons operably-linked to the transposase and/or to the gene of interest. In an alternate embodiment, one stop codon of UAA or UGA is operably linked to the transposase and/or to the gene of interest. As used herein an "effective polyA sequence" refers to either a synthetic or non-synthetic sequence that contains multiple and sequential nucleotides containing an adenine base (an A polynucleotide string) and that increases expression of the gene to which it is operably-linked. A polyA sequence may be operably-linked to any gene in the transposon-based vector including, but not limited to, a transposase gene and a gene of interest. In one embodiment, a polyA sequence comprises the polynucleotide sequence provided in SEQ ID NO:28. A preferred polyA sequence is optimized for use in the host animal or human. In one embodiment, the polyA sequence is optimized for use in a bird, and more specifically, a chicken. The chicken optimized polyA sequence generally contains a minimum of 60 base pairs, and more preferably between approximately 60 and several hundred base pairs, that precede the A polynucleotide string and thereby separate the stop codon from the A polynucleotide string. A chicken optimized polyA sequence may also have a reduced amount of CT repeats as compared to a synthetic polyA sequence. In one embodiment of the present invention, the polyA sequence comprises a conalbumin polyA sequence as provided in SEQ ID NO:33 and as taken from GenBank accession # Y00407, base pairs 10651-11058.

Promoters and Enhancers

The first promoter operably-linked to the transposase gene and the second promoter operably-linked to the gene of interest can be a constitutive promoter or an inducible promoter. Constitutive promoters include, but are not limited to, immediate 5 early cytomegalovirus (CMV) promoter, herpes simplex virus 1 (HSV1) immediate early promoter, SV40 promoter, lysozyme promoter, early and late CMV promoters, early and late HSV promoters, β -actin promoter, tubulin promoter, Rous-Sarcoma virus (RSV) promoter, and heat-shock protein (HSP) promoter. Inducible promoters include tissue-specific promoters, developmentally-regulated promoters and 10 chemically inducible promoters. Examples of tissue-specific promoters include the glucose 6 phosphate (G6P) promoter, vitellogenin promoter, ovalbumin promoter, ovomucoid promoter, conalbumin promoter, ovotransferrin promoter, prolactin promoter, kidney uromodulin promoter, and placental lactogen promoter. In one embodiment, the vitellogenin promoter includes a polynucleotide sequence of SEQ ID 15 NO:17. The G6P promoter sequence may be deduced from a rat G6P gene untranslated upstream region provided in GenBank accession number U57552.1. Examples of developmentally-regulated promoters include the homeobox promoters and several hormone induced promoters. Examples of chemically inducible promoters include reproductive hormone induced promoters and antibiotic inducible 20 promoters such as the tetracycline inducible promoter and the zinc-inducible metallothionein promoter.

Other inducible promoter systems include the Lac operator repressor system inducible by IPTG (isopropyl beta-D-thiogalactoside) (Cronin, A. et al. 2001. Genes and Development, v. 15), ecdysone-based inducible systems (Hoppe, U. C. et al. 25 2000. Mol. Ther. 1:159-164); estrogen-based inducible systems (Braselmann, S. et al. 1993. Proc. Natl. Acad. Sci. 90:1657-1661); progesterone-based inducible systems using a chimeric regulator, GLVP, which is a hybrid protein consisting of the GAL4 binding domain and the herpes simplex virus transcriptional activation domain, VP16, and a truncated form of the human progesterone receptor that retains the ability to 30 bind ligand and can be turned on by RU486 (Wang, et al. 1994. Proc. Natl. Acad. Sci. 91:8180-8184); CID-based inducible systems using chemical inducers of dimerization (CIDs) to regulate gene expression, such as a system wherein rapamycin induces dimerization of the cellular proteins FKBP12 and FRAP (Belshaw, P. J. et al. 1996. J.

Chem. Biol. 3:731-738; Fan, L. et al. 1999. Hum. Gene Ther. 10:2273-2285; Shariat, S.F. et al. 2001. Cancer Res. 61:2562-2571; Spencer, D.M. 1996. Curr. Biol. 6:839-847). Chemical substances that activate the chemically inducible promoters can be administered to the animal containing the transgene of interest via any method known to those of skill in the art.

5 Other examples of cell or tissue-specific and constitutive promoters include but are not limited to smooth-muscle SM22 promoter, including chimeric SM22alpha/telokin promoters (Hoggatt A.M. et al., 2002. Circ Res. 91(12):1151-9); ubiquitin C promoter (Biochim Biophys Acta, 2003. Jan. 3;1625(1):52-63); Hsf2 10 promoter; murine COMP (cartilage oligomeric matrix protein) promoter; early B cell-specific mb-1 promoter (Sigvardsson M., et al., 2002. Mol. Cell Biol. 22(24):8539-51); prostate specific antigen (PSA) promoter (Yoshimura I. et al., 2002, J. Urol. 168(6):2659-64); exorh promoter and pineal expression-promoting element (Asaoka 15 Y., et al., 2002. Proc. Natl. Acad. Sci. 99(24):15456-61); neural and liver ceramidase gene promoters (Okino N. et al., 2002. Biochem. Biophys. Res. Commun. 299(1):160-6); PSP94 gene promoter/enhancer (Gabril M.Y. et al., 2002. Gene Ther. 9(23):1589-99); promoter of the human FAT/CD36 gene (Kuriki C., et al., 2002. Biol. Pharm. Bull. 25(11):1476-8); VL30 promoter (Staplin W.R. et al., 2002. Blood 20 October 24, 2002); IL-10 promoter (Brenner S., et al., 2002. J. Biol. Chem. December 20 18, 2002).

Examples of avian promoters include, but are not limited to, promoters controlling expression of egg white proteins, such as ovalbumin, ovotransferrin (conalbumin), ovomucoid, lysozyme, ovomucin, g2 ovoglobulin, g3 ovoglobulin, ovoflavoprotein, ovostatin (ovomacroglobin), cystatin, avidin, thiamine-binding 25 protein, glutamyl aminopeptidase minor glycoprotein 1, minor glycoprotein 2; and promoters controlling expression of egg-yolk proteins, such as vitellogenin, very low-density lipoproteins, low density lipoprotein, cobalamin-binding protein, riboflavin-binding protein, biotin-binding protein (Awade, 1996. Z. Lebensm. Unters. Forsch. 202:1-14). An advantage of using the vitellogenin promoter is that it is active during 30 the egg-laying stage of an animal's life-cycle, which allows for the production of the protein of interest to be temporally connected to the import of the protein of interest into the egg yolk when the protein of interest is equipped with an appropriate targeting sequence.

Liver-specific promoters of the present invention include, but are not limited to, the following promoters, vitellogenin promoter, G6P promoter, cholesterol-7-alpha-hydroxylase (CYP7A) promoter, phenylalanine hydroxylase (PAH) promoter, protein C gene promoter, insulin-like growth factor I (IGF-I) promoter, bilirubin 5 UDP-glucuronosyltransferase promoter, aldolase B promoter, furin promoter, metallothioneine promoter, albumin promoter, and insulin promoter.

Also included in the present invention are promoters that can be used to target expression of a protein of interest into the milk of a milk-producing animal including, but not limited to, β lactoglobin promoter, whey acidic protein promoter, lactalbumin 10 promoter and casein promoter.

Promoters associated with cells of the immune system may also be used. Acute phase promoters such as interleukin (IL)-1 and IL-2 may be employed. Promoters for heavy and light chain Ig may also be employed. The promoters of the T cell receptor components CD4 and CD8, B cell promoters and the promoters of 15 CR2 (complement receptor type 2) may also be employed. Immune system promoters are preferably used when the desired protein is an antibody protein.

Also included in this invention are modified promoters/enhancers wherein elements of a single promoter are duplicated, modified, or otherwise changed. In one embodiment, steroid hormone-binding domains of the ovalbumin promoter are moved 20 from about -6.5 kb to within approximately the first 1000 base pairs of the gene of interest. Modifying an existing promoter with promoter/enhancer elements not found naturally in the promoter, as well as building an entirely synthetic promoter, or drawing promoter/enhancer elements from various genes together on a non-natural backbone, are all encompassed by the current invention.

25 Accordingly, it is to be understood that the promoters contained within the transposon-based vectors of the present invention may be entire promoter sequences or fragments of promoter sequences. For example, in one embodiment, the promoter operably linked to a gene of interest is an approximately 900 base pair fragment of a chicken ovalbumin promoter (SEQ ID NO:40). The constitutive and inducible 30 promoters contained within the transposon-based vectors may also be modified by the addition of one or more modified Kozak sequences of ACCATG (SEQ ID NO:13).

As indicated above, the present invention includes transposon-based vectors containing one or more enhancers. These enhancers may or may not be operably-

linked to their native promoter and may be located at any distance from their operably-linked promoter. A promoter operably-linked to an enhancer is referred to herein as an "enhanced promoter." The enhancers contained within the transposon-based vectors are preferably enhancers found in birds, and more preferably, an 5 ovalbumin enhancer, but are not limited to these types of enhancers. In one embodiment, an approximately 675 base pair enhancer element of an ovalbumin promoter is cloned upstream of an ovalbumin promoter with 300 base pairs of spacer DNA separating the enhancer and promoter. In one embodiment, the enhancer used as a part of the present invention comprises base pairs 1-675 of a Chicken Ovalbumin 10 enhancer from GenBank accession #S82527.1. The polynucleotide sequence of this enhancer is provided in SEQ ID NO:37.

Also included in some of the transposon-based vectors of the present invention are cap sites and fragments of cap sites. In one embodiment, approximately 50 base pairs of a 5' untranslated region wherein the capsites resides are added on the 3' end of 15 an enhanced promoter or promoter. An exemplary 5' untranslated region is provided in SEQ ID NO:38. A putative cap-site residing in this 5' untranslated region preferably comprises the polynucleotide sequence provided in SEQ ID NO:39.

In one embodiment of the present invention, the first promoter operably-linked to the transposase gene is a constitutive promoter and the second promoter operably-linked to the gene of interest is a tissue-specific promoter. In this embodiment, use of 20 the first constitutive promoter allows for constitutive activation of the transposase gene and incorporation of the gene of interest into virtually all cell types, including the germline of the recipient animal. Although the gene of interest is incorporated into the germline generally, the gene of interest is only expressed in a tissue-specific 25 manner. It should be noted that cell- or tissue-specific expression as described herein does not require a complete absence of expression in cells or tissues other than the preferred cell or tissue. Instead, "cell-specific" or "tissue-specific" expression refers to a majority of the expression of a particular gene of interest in the preferred cell or tissue, respectively.

30 When incorporation of the gene of interest into the germline is not preferred, the first promoter operably-linked to the transposase gene can be a tissue-specific promoter. For example, transfection of a transposon-based vector containing a transposase gene operably-linked to a liver-specific promoter such as the G6P

promoter or vitellogenin promoter provides for activation of the transposase gene and incorporation of the gene of interest in the cells of the liver but not into the germline and other cells generally. In this second embodiment, the second promoter operably-linked to the gene of interest can be a constitutive promoter or an inducible promoter.

5 In a preferred embodiment, both the first promoter and the second promoter are a G6P promoter. In embodiments wherein tissue-specific expression or incorporation is desired, it is preferred that the transposon-based vector is administered directly to the tissue of interest or to an artery leading to the tissue of interest.

Accordingly, cell specific promoters may be used to enhance transcription in selected tissues. In birds, for example, promoters that are found in cells of the fallopian tube, such as ovalbumin, conalbumin, ovomucoid and/or lysozyme, are used in the vectors to ensure transcription of the gene of interest in the epithelial cells and tubular gland cells of the fallopian tube, leading to synthesis of the desired protein encoded by the gene and deposition into the egg white. In mammals, promoters specific for the epithelial cells of the alveoli of the mammary gland, such as prolactin, insulin, beta lactoglobin, whey acidic protein, lactalbumin, casein, and/or placental lactogen, are used in the design of vectors used for transfection of these cells for the production of desired proteins for deposition into the milk. In liver cells, the G6P promoter may be employed to drive transcription of the gene of interest for protein production. Proteins made in the liver of birds may be delivered to the egg yolk.

In order to achieve higher or more efficient expression of the transposase gene, the promoter and other regulatory sequences operably-linked to the transposase gene may be those derived from the host. These host specific regulatory sequences can be tissue specific as described above or can be of a constitutive nature. For example, an avian actin promoter and its associated polyA sequence can be operably-linked to a transposase in a transposase-based vector for transfection into an avian. Examples of other host specific promoters that could be operably-linked to the transposase include the myosin and DNA or RNA polymerase promoters.

Directing Sequences

30 In some embodiments of the present invention, the gene of interest is operably-linked to a directing sequence or a sequence that provides proper conformation to the desired protein encoded by the gene of interest. As used herein, the term "directing sequence" refers to both signal sequences and targeting sequences.

An egg directing sequence includes, but is not limited to, an ovomucoid signal sequence, an ovalbumin signal sequence and a vitellogenin targeting sequence. The term "signal sequence" refers to an amino acid sequence, or the polynucleotide sequence that encodes the amino acid sequence, that directs the protein to which it is linked to the endoplasmic reticulum in a eukaryote, and more preferably the translocational pores in the endoplasmic reticulum, or the plasma membrane in a prokaryote, or mitochondria, such as for the purpose of gene therapy of mitochondrial diseases. Signal and targeting sequences can be used to direct a desired protein into, for example, the milk, when the transposon-based vectors are administered to a milk-producing animal.

Signal sequences can also be used to direct a desired protein into, for example, a secretory pathway for incorporation into the egg yolk or the egg white, when the transposon-based vectors are administered to a bird or other egg-laying animal. One example of such a transposon-based vector is provided in Figure 3 wherein the gene of interest is operably linked to the ovomucoid signal sequence. The present invention also includes a gene of interest operably-linked to a second gene containing a signal sequence. An example of such an embodiment is shown in Figure 2 wherein the gene of interest is operably-linked to the ovalbumin gene that contains an ovalbumin signal sequence. Other signal sequences that can be included in the transposon-based vectors include, but are not limited to the ovotransferrin and lysozyme signal sequences.

As also used herein, the term "targeting sequence" refers to an amino acid sequence, or the polynucleotide sequence encoding the amino acid sequence, which amino acid sequence is recognized by a receptor located on the exterior of a cell. Binding of the receptor to the targeting sequence results in uptake of the protein or peptide operably-linked to the targeting sequence by the cell. One example of a targeting sequence is a vitellogenin targeting sequence that is recognized by a vitellogenin receptor (or the low density lipoprotein receptor) on the exterior of an oocyte. In one embodiment, the vitellogenin targeting sequence includes the polynucleotide sequence of SEQ ID NO:18. In another embodiment, the vitellogenin targeting sequence includes all or part of the vitellogenin gene. Other targeting sequences include VLDL and Apo E, which are also capable of binding the vitellogenin receptor. Since the ApoE protein is not endogenously expressed in birds,

its presence may be used advantageously to identify birds carrying the transposon-based vectors of the present invention.

Genes of Interest Encoding Desired Proteins

A gene of interest selected for stable incorporation is designed to encode any 5 desired protein or peptide or to regulate any cellular response. In some embodiments, the desired proteins or peptides are deposited in an egg or in milk. It is to be understood that the present invention encompasses transposon-based vectors containing multiple genes of interest. The multiple genes of interest may each be operably-linked to a separate promoter and other regulatory sequence(s) or may all be 10 operably-linked to the same promoter and other regulatory sequences(s). In one embodiment, multiple gene of interest are linked to a single promoter and other regulatory sequence(s) and each gene of interest is separated by a cleavage site or a pro portion of a signal sequence.

Protein and peptide hormones are a preferred class of proteins in the present 15 invention. Such protein and peptide hormones are synthesized throughout the endocrine system and include, but are not limited to, hypothalamic hormones and hypophysiotropic hormones, anterior, intermediate and posterior pituitary hormones, pancreatic islet hormones, hormones made in the gastrointestinal system, renal hormones, thymic hormones, parathyroid hormones, adrenal cortical and medullary 20 hormones. Specifically, hormones that can be produced using the present invention include, but are not limited to, chorionic gonadotropin, corticotropin, erythropoietin, glucagons, IGF-1, oxytocin, platelet-derived growth factor, calcitonin, follicle-stimulating hormone, leutinizing hormone, thyroid-stimulating hormone, insulin, gonadotropin-releasing hormone and its analogs, vasopressin, octreotide, 25 somatostatin, prolactin, adrenocorticotrophic hormone, antidiuretic hormone, thyrotropin-releasing hormone (TRH), growth hormone-releasing hormone (GHRH), dopamine, melatonin, thyroxin (T₄), parathyroid hormone (PTH), glucocorticoids such as cortisol, mineralocorticoids such as aldosterone, androgens such as testosterone, adrenaline (epinephrine), noradrenaline (norepinephrine), estrogens such 30 as estradiol, progesterone, glucagons, calcitonin, calciferol, atrial-natriuretic peptide, gastrin, secretin, cholecystokinin (CCK), neuropeptide Y, ghrelin, PYY₃₋₃₆, angiotensinogen, thrombopoietin, and leptin. By using appropriate polynucleotide sequences, species-specific hormones may be made by transgenic animals.

In one embodiment of the present invention, the gene of interest is a proinsulin gene and the desired molecule is insulin. Proinsulin consists of three parts: a C-peptide and two long strands of amino acids (called the alpha and beta chains) that later become linked together to form the insulin molecule. Figures 2 and 3 are 5 schematics of transposon-based vector constructs containing a proinsulin gene operably-linked to an ovalbumin promoter and ovalbumin protein or an ovomucoid promoter and ovomucoid signal sequence, respectively. In these embodiments, proinsulin is expressed in the oviduct tubular gland cells and then deposited in the egg white. One example of a proinsulin polynucleotide sequence is shown in SEQ ID 10 NO:21, wherein the C-peptide cleavage site spans from Arg at position 31 to Arg at position 65.

Serum proteins including lipoproteins such as high density lipoprotein (HDL), HDL-Milano and low density lipoprotein, albumin, clotting cascade factors, factor 15 VIII, factor IX, fibrinogen, and globulins are also included in the group of desired proteins of the present invention. Immunoglobulins are one class of desired globulin molecules and include but are not limited to IgG, IgM, IgA, IgD, IgE, IgY, lambda chains, kappa chains and fragments thereof; Fc fragments, and Fab fragments. Desired antibodies include, but are not limited to, naturally occurring antibodies, 20 human antibodies, humanized antibodies, and hybrid antibodies. Genes encoding modified versions of naturally occurring antibodies or fragments thereof and genes 25 encoding artificially designed antibodies or fragments thereof may be incorporated into the transposon-based vectors of the present invention. Desired antibodies also include antibodies with the ability to bind specific ligands, for example, antibodies against proteins associated with cancer-related molecules, such as anti-her 2, or anti-CA125. Accordingly, the present invention encompasses a transposon-based vector 30 containing one or more genes encoding a heavy immunoglobulin (Ig) chain and a light Ig chain. Further, more than one gene encoding for more than one antibody may be administered in one or more transposon-based vectors of the present invention. In this manner, an egg may contain more than one type of antibody in the egg white, the egg yolk or both.

In one embodiment, a transposon-based vector contains a heavy Ig chain and a light Ig chain, both operably linked to a promoter. Figures 5 and 6 schematically depict exemplary constructs of this embodiment. More specifically, Figure 5 shows a

construct containing a cecropin pre-pro sequence and a cecropin pro sequence, wherein the pre sequence functions to direct the resultant protein into the endoplasmic reticulum and the pro sequences and the pro sequences are cleaved upon secretion of the protein from a cell into which the construct has been transfected. Figure 6 shows 5 a construct containing an enterokinase cleavage site. In this embodiment, it may be required to further remove several additional amino acids from the light chain following cleavage by enterokinase. In another embodiment, the transposon-based vector comprises a heavy Ig chain operably-linked to one promoter and a light Ig chain operably-linked to another promoter. Figure 7 schematically depicts an 10 exemplary construct of this embodiment. The present invention also encompasses a transposon-based vector containing genes encoding portions of a heavy Ig chain and/or portions of a light Ig chain. The present invention further includes a transposon-based vector containing a gene that encodes a fusion protein comprising a heavy and/or light Ig chain, or portions thereof.

15 Antibodies used as therapeutic reagents include but are not limited to antibodies for use in cancer immunotherapy against specific antigens, or for providing passive immunity to an animal or a human against an infectious disease or a toxic agent. Antibodies used as diagnostic reagents include, but are not limited to antibodies that may be labeled and detected with a detector, for example antibodies 20 with a fluorescent label attached that may be detected following exposure to specific wavelengths. Such labeled antibodies may be primary antibodies directed to a specific antigen, for example, rhodamine-labeled rabbit anti-growth hormone, or may be labeled secondary antibodies, such as fluorescein-labeled goat-anti chicken IgG. Such labeled antibodies are known to one of ordinary skill in the art. Labels useful 25 for attachment to antibodies are also known to one of ordinary skill in the art. Some of these labels are described in the "Handbook of Fluorescent Probes and Research Products", ninth edition, Richard P. Haugland (ed) Molecular Probes, Inc. Eugene, OR), which is incorporated herein in its entirety.

Antibodies produced with using the present invention may be used as 30 laboratory reagents for numerous applications including radioimmunoassay, western blots, dot blots, ELISA, immunoaffinity columns and other procedures requiring antibodies as known to one of ordinary skill in the art. Such antibodies include

primary antibodies, secondary antibodies and tertiary antibodies, which may be labeled or unlabeled.

Antibodies that may be made with the practice of the present invention include, but are not limited to primary antibodies, secondary antibodies, designer 5 antibodies, anti-protein antibodies, anti-peptide antibodies, anti-DNA antibodies, anti-RNA antibodies, anti-hormone antibodies, anti-hypophysiotropic peptides, antibodies against non-natural antigens, anti-anterior pituitary hormone antibodies, anti-posterior pituitary hormone antibodies, anti-venom antibodies, anti-tumor marker antibodies, antibodies directed against epitopes associated with infectious disease, including, anti- 10 viral, anti-bacterial, antiprotozoal, anti-fungal, anti-parasitic, anti-receptor, anti-lipid, anti-phospholipid, anti-growth factor, anti-cytokine, anti-monokine, anti-idiotype, and anti-accessory (presentation) protein antibodies. Antibodies made with the present invention, as well as light chains or heavy chains, may also be used to inhibit enzyme activity.

15 Antibodies that may be produced using the present invention include, but are not limited to, antibodies made against the following proteins: Bovine γ -Globulin, Serum; Bovine IgG, Plasma; Chicken γ -Globulin, Serum; Human γ -Globulin, Serum; Human IgA, Plasma; Human IgA₁, Myeloma; Human IgA₂, Myeloma; Human IgA₂, Plasma; Human IgD, Plasma; Human IgE, Myeloma; Human IgG, Plasma; Human 20 IgG, Fab Fragment, Plasma; Human IgG, F(ab')₂ Fragment, Plasma; Human IgG, Fc Fragment, Plasma; Human IgG₁, Myeloma; Human IgG₂, Myeloma; Human IgG₃, Myeloma; Human IgG₄, Myeloma; Human IgM, Myeloma; Human IgM, Plasma; Human Immunoglobulin, Light Chain κ , Urine; Human Immunoglobulin, Light 25 Chains κ and λ , Plasma; Mouse γ -Globulin, Serum; Mouse IgG, Serum; Mouse IgM, Myeloma; Rabbit γ -Globulin, Serum; Rabbit IgG, Plasma; and Rat γ -Globulin, Serum. In one embodiment, the transposon-based vector comprises the coding sequence of light and heavy chains of a murine monoclonal antibody that shows specificity for human seminoprotein (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively).

30 A further non-limiting list of antibodies that recognize other antibodies is as follows: Anti-Chicken IgG, heavy (H) & light (L) Chain Specific (Sheep); Anti-Goat γ -Globulin (Donkey); Anti-Goat IgG, Fc Fragment Specific (Rabbit); Anti-Guinea Pig γ -Globulin (Goat); Anti-Human Ig, Light Chain, Type κ Specific; Anti-Human Ig,

Light Chain, Type λ Specific; Anti-Human IgA, α -Chain Specific (Goat); Anti-Human IgA, Fab Fragment Specific; Anti-Human IgA, Fc Fragment Specific; Anti-Human IgA, Secretory; Anti-Human IgE, ϵ -Chain Specific (Goat); Anti-Human IgE, Fc Fragment Specific; Anti-Human IgG, Fc Fragment Specific (Goat); Anti-Human IgG, γ -Chain Specific (Goat); Anti-Human IgG, Fc Fragment Specific; Anti-Human IgG, Fd Fragment Specific; Anti-Human IgG, H & L Chain Specific (Goat); Anti-Human IgG₁, Fc Fragment Specific; Anti-Human IgG₂, Fc Fragment Specific; Anti-Human IgG₂, Fd Fragment Specific; Anti-Human IgG₃, Hinge Specific; Anti-Human IgG₄, Fc Fragment Specific; Anti-Human IgM, Fc Fragment Specific; Anti-Human IgM, μ -Chain Specific; Anti-Mouse IgE, ϵ -Chain Specific; Anti-Mouse γ -Globulin (Goat); Anti-Mouse IgG, γ -Chain Specific (Goat); Anti-Mouse IgG, H & L Chain Specific (Goat); Anti-Mouse IgM, μ -Chain Specific (Goat); Anti-Mouse IgM, H & L Chain Specific (Goat); Anti-Rabbit γ -Globulin (Goat); Anti-Rabbit IgG, Fc Fragment Specific (Goat); Anti-Rabbit IgG, H & L Chain Specific (Goat); Anti-Rat γ -Globulin (Goat); Anti-Rat IgG, H & L Chain Specific; Anti-Rhesus Monkey γ -Globulin (Goat); and, Anti-Sheep IgG, H & L Chain Specific.

Another non-limiting list of the antibodies that may be produced using the present invention is provided in product catalogs of companies such as Phoenix 20 Pharmaceuticals, Inc. (www.phoenixpeptide.com; 530 Harbor Boulevard, Belmont, CA), Peninsula Labs San Carlos CA, SIGMA, St.Louis, MO www.sigmaldrich.com, Cappel ICN, Irvine, California, www.icnbiomed.com, and Calbiochem, La Jolla, California, www.calbiochem.com, which are all incorporated herein by reference in their entirety. The polynucleotide sequences encoding these antibodies 25 may be obtained from the scientific literature, from patents, and from databases such as GenBank. Alternatively, one of ordinary skill in the art may design the polynucleotide sequence to be incorporated into the genome by choosing the codons that encode for each amino acid in the desired antibody. Antibodies made by the transgenic animals of the present invention include antibodies that may be used as 30 therapeutic reagents, for example in cancer immunotherapy against specific antigens, as diagnostic reagents and as laboratory reagents for numerous applications including immunoneutralization, radioimmunoassay, western blots, dot blots, ELISA, immunoprecipitation and immunoaffinity columns. Some of these antibodies include,

but are not limited to, antibodies which bind the following ligands: adrenomedulin, amylin, calcitonin, amyloid, calcitonin gene-related peptide, cholecystokinin, gastrin, gastric inhibitory peptide, gastrin releasing peptide, interleukin, interferon, cortistatin, somatostatin, endothelin, sarafotoxin, glucagon, glucagon-like peptide, insulin, atrial 5 natriuretic peptide, BNP, CNP, neurokinin, substance P, leptin, neuropeptide Y, melanin concentrating hormone, melanocyte stimulating hormone, orphanin, endorphin, dynorphin, enkephalin, enkephalin, leumorphin, peptide F, PACAP, PACAP-related peptide, parathyroid hormone, urocortin, corticotrophin releasing hormone, PHM, PHI, vasoactive intestinal polypeptide, secretin, ACTH, angiotensin, 10 angiotatin, bombesin, endostatin, bradykinin, FMRF amide, galanin, gonadotropin releasing hormone (GnRH) associated peptide, GnRH, growth hormone releasing hormone, inhibin, granulocyte-macrophage colony stimulating factor (GM-CSF), motilin, neurotensin, oxytocin, vasopressin, osteocalcin, pancreastatin, pancreatic polypeptide, peptide YY, proopiomelanocortin, transforming growth factor, vascular 15 endothelial growth factor, vesicular monoamine transporter, vesicular acetylcholine transporter, ghrelin, NPW, NPB, C3d, prokinetican, thyroid stimulating hormone, luteinizing hormone, follicle stimulating hormone, prolactin, growth hormone, beta-lipotropin, melatonin, kallikriens, kinins, prostaglandins, erythropoietin, p146 (SEQ ID NO:18 amino acid sequence, SEQ ID NO:19, nucleotide sequence), estrogen, 20 testosterone, corticosteroids, mineralocorticoids, thyroid hormone, thymic hormones, connective tissue proteins, nuclear proteins, actin, avidin, activin, agrin, albumin, and prohormones, propeptides, splice variants, fragments and analogs thereof.

The following is yet another non-limiting of antibodies that can be produced by the methods of present invention: abciximab (ReoPro), abciximab anti-platelet 25 aggregation monoclonal antibody, anti-CD11a (hu1124), anti-CD18 antibody, anti-CD20 antibody, anti-cytomegalovirus (CMV) antibody, anti-digoxin antibody, anti-hepatitis B antibody, anti-HER-2 antibody, anti-idiotype antibody to GD3 glycolipid, anti-IgE antibody, anti-IL-2R antibody, antimetastatic cancer antibody (mAb 17-1A), anti-rabies antibody, anti-respiratory syncytial virus (RSV) antibody, anti-Rh 30 antibody, anti-TCR, anti-TNF antibody, anti-VEGF antibody and fab fragment thereof, rattlesnake venom antibody, black widow spider venom antibody, coral snake venom antibody, antibody against very late antigen-4 (VLA-4), C225 humanized antibody to EGF receptor, chimeric (human & mouse) antibody against TNF α ,

antibody directed against GPIb/IIIa receptor on human platelets, gamma globulin, anti-hepatitis B immunoglobulin, human anti-D immunoglobulin, human antibodies against *S aureus*, human tetanus immunoglobulin, humanized antibody against the epidermal growth receptor-2, humanized antibody against the α subunit of the 5 interleukin-2 receptor, humanized antibody CTLA4IG, humanized antibody to the IL-2 R α -chain, humanized anti-CD40-ligand monoclonal antibody (5c8), humanized mAb against the epidermal growth receptor-2, humanized mAb to rous sarcoma virus, humanized recombinant antibody (IgG1k) against respiratory syncytial virus (RSV), lymphocyte immunoglobulin (anti-thymocyte antibody), lymphocyte 10 immunoglobulin, mAb against factor VII, MDX-210 bi-specific antibody against HER-2, MDX-22, MDX-220 bi-specific antibody against TAG-72 on tumors, MDX-33 antibody to Fc γ R1 receptor, MDX-447 bi-specific antibody against EGF receptor, MDX-447 bispecific humanized antibody to EGF receptor, MDX-RA immunotoxin (ricin A linked) antibody, Medi-507 antibody (humanized form of BTI-322) against 15 CD2 receptor on T-cells, monoclonal antibody LDP-02, muromonab-CD3(OKT3) antibody, OKT3 ("muromonab-CD3") antibody, PRO 542 antibody, ReoPro ("abciximab") antibody, and TNF-IgG fusion protein.

The antibodies prepared using the methods of the present invention may also be designed to possess specific labels that may be detected through means known to 20 one of ordinary skill in the art. The antibodies may also be designed to possess specific sequences useful for purification through means known to one of ordinary skill in the art. Specialty antibodies designed for binding specific antigens may also be made in transgenic animals using the transposon-based vectors of the present invention.

25 Production of a monoclonal antibody using the transposon-based vectors of the present invention can be accomplished in a variety of ways. In one embodiment, two vectors may be constructed: one that encodes the light chain, and a second vector that encodes the heavy chain of the monoclonal antibody. These vectors may then be incorporated into the genome of the target animal by methods disclosed herein. In an 30 alternative embodiment, the sequences encoding light and heavy chains of a monoclonal antibody may be included on a single DNA construct. For example, the coding sequence of light and heavy chains of a murine monoclonal antibody that show specificity for human seminoprotein can be expressed using transposon-based

constructs of the present invention (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively).

Further included in the present invention are proteins and peptides synthesized by the immune system including those synthesized by the thymus, lymph nodes, spleen, and the gastrointestinal associated lymph tissues (GALT) system.

5 The immune system proteins and peptides proteins that can be made in transgenic animals using the transposon-based vectors of the present invention include, but are not limited to, alpha-interferon, beta-interferon, gamma-interferon, alpha-interferon A, alpha-interferon 1, G-CSF, GM-CSF, interleukin-1 (IL-1), IL-2, IL-3, IL-4, IL-5, IL-6, 10 IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, TNF- α , and TNF- β . Other cytokines included in the present invention include cardiotrophin, stromal cell derived factor, macrophage derived chemokine (MDC), melanoma growth stimulatory activity (MGSA), macrophage inflammatory proteins 1 alpha (MIP-1 alpha), 2, 3 alpha, 3 beta, 4 and 5.

15 Lytic peptides such as p146 are also included in the desired molecules of the present invention. In one embodiment, the p146 peptide comprises an amino acid sequence of SEQ ID NO:19. The present invention also encompasses a transposon-based vector comprising a p146 nucleic acid comprising a polynucleotide sequence of SEQ ID NO:20.

20 Enzymes are another class of proteins that may be made through the use of the transposon-based vectors of the present invention. Such enzymes include but are not limited to adenosine deaminase, alpha-galactosidase, cellulase, collagenase, dnaseI, hyaluronidase, lactase, L-asparaginase, pancreatin, papain, streptokinase B, subtilisin, superoxide dismutase, thrombin, trypsin, urokinase, fibrinolysin, glucocerebrosidase

25 and plasminogen activator. In some embodiments wherein the enzyme could have deleterious effects, additional amino acids and a protease cleavage site are added to the carboxy end of the enzyme of interest in order to prevent expression of a functional enzyme. Subsequent digestion of the enzyme with a protease results in activation of the enzyme.

30 Extracellular matrix proteins are one class of desired proteins that may be made through the use of the present invention. Examples include but are not limited to collagen, fibrin, elastin, laminin, and fibronectin and subtypes thereof. Intracellular

proteins and structural proteins are other classes of desired proteins in the present invention.

Growth factors are another desired class of proteins that may be made through the use of the present invention and include, but are not limited to, transforming 5 growth factor- α ("TGF- α "), transforming growth factor- β (TGF- β), platelet-derived growth factors (PDGF), fibroblast growth factors (FGF), including FGF acidic isoforms 1 and 2, FGF basic form 2 and FGF 4, 8, 9 and 10, nerve growth factors (NGF) including NGF 2.5s, NGF 7.0s and beta NGF and neurotrophins, brain derived neurotrophic factor, cartilage derived factor, growth factors for stimulation of the 10 production of red blood cells, growth factors for stimulation of the production of white blood cells, bone growth factors (BGF), basic fibroblast growth factor, vascular endothelial growth factor (VEGF), granulocyte colony stimulating factor (G-CSF), insulin like growth factor (IGF) I and II, hepatocyte growth factor, glial neurotrophic growth factor (GDNF), stem cell factor (SCF), keratinocyte growth factor (KGF), 15 transforming growth factors (TGF), including TGFs alpha, beta, beta1, beta2, beta3, skeletal growth factor, bone matrix derived growth factors, bone derived growth factors, erythropoietin (EPO) and mixtures thereof.

Another desired class of proteins that may be made may be made through the use of the present invention include but are not limited to leptin, leukemia inhibitory 20 factor (LIF), tumor necrosis factor alpha and beta, ENBREL, angiostatin, endostatin, thrombospondin, osteogenic protein-1, bone morphogenetic proteins 2 and 7, osteonectin, somatomedin-like peptide, and osteocalcin.

A non-limiting list of the peptides and proteins that may be made may be made through the use of the present invention is provided in product catalogs of 25 companies such as Phoenix Pharmaceuticals, Inc. (www.phoenixpeptide.com; 530 Harbor Boulevard • Belmont, CA), Peninsula Labs San Carlos CA, SIGMA, St.Louis, MO www.sigma-aldrich.com, Cappel ICN, Irvine, California, www.icnbiomed.com, and Calbiochem, La Jolla, California, www.calbiochem.com. The polynucleotide sequences encoding these proteins and peptides of interest may be obtained from the 30 scientific literature, from patents, and from databases such as GenBank. Alternatively, one of ordinary skill in the art may design the polynucleotide sequence to be incorporated into the genome by choosing the codons that encode for each amino acid in the desired protein or peptide.

Some of these desired proteins or peptides that may be made through the use of the present invention include but are not limited to the following: adrenomedulin, amylin, calcitonin, amyloid, calcitonin gene-related peptide, cholecystokinin, gastrin, gastric inhibitory peptide, gastrin releasing peptide, interleukin, interferon, cortistatin, 5 somatostatin, endothelin, sarafotoxin, glucagon, glucagon-like peptide, insulin, atrial natriuretic peptide, BNP, CNP, neurokinin, substance P, leptin, neuropeptide Y, melanin concentrating hormone, melanocyte stimulating hormone, orphanin, endorphin, dynorphin, enkephalin, leumorphin, peptide F, PACAP, PACAP-related peptide, parathyroid hormone, urocortin, corticotrophin releasing hormone, PHM, 10 PHI, vasoactive intestinal polypeptide, secretin, ACTH, angiotensin, angiotatin, bombesin, endostatin, bradykinin, FMRF amide, galanin, gonadotropin releasing hormone (GnRH) associated peptide, GnRH, growth hormone releasing hormone, inhibin, granulocyte-macrophage colony stimulating factor (GM-CSF), motilin, neurotensin, oxytocin, vasopressin, osteocalcin, pancreastatin, pancreatic polypeptide, 15 peptide YY, proopiomelanocortin, transforming growth factor, vascular endothelial growth factor, vesicular monoamine transporter, vesicular acetylcholine transporter, ghrelin, NPW, NPB, C3d, prokinetican, thyroid stimulating hormone, luteinizing hormone, follicle stimulating hormone, prolactin, growth hormone, beta-lipotropin, melatonin, kallikriens, kinins, prostaglandins, erythropoietin, p146 (SEQ ID NO:19, 20 amino acid sequence, SEQ ID NO:20, nucleotide sequence), thymic hormones, connective tissue proteins, nuclear proteins, actin, avidin, activin, agrin, albumin, and prohormones, propeptides, splice variants, fragments and analogs thereof.

Other desired proteins that may be made by the transgenic animals of the present invention include bacitracin, polymixin b, vancomycin, cyclosporine, anti- 25 RSV antibody, alpha-1 antitrypsin (AAT), anti-cytomegalovirus antibody, anti-hepatitis antibody, anti-inhibitor coagulant complex, anti-rabies antibody, anti-Rh(D) antibody, adenosine deaminase, anti-digoxin antibody, antivenin crotalidae (rattlesnake venom antibody), antivenin latrodectus (black widow spider venom antibody), antivenin micrurus (coral snake venom antibody), aprotinin, corticotropin 30 (ACTH), diphtheria antitoxin, lymphocyte immune globulin (anti-thymocyte antibody), protamine, thyrotropin, capreomycin, α -galactosidase, gramicidin, streptokinase, tetanus toxoid, tyrothricin, IGF-1, proteins of varicella vaccine, anti-TNF antibody, anti-IL-2r antibody, anti-HER-2 antibody, OKT3 ("muromonab-

CD3") antibody, TNF-IgG fusion protein, ReoPro ("abciximab") antibody, ACTH fragment 1-24, desmopressin, gonadotropin-releasing hormone, histrelin, leuprolide, lypressin, nafarelin, peptide that binds GPIIb/GPIIIa on platelets (integrilin), goserelin, capreomycin, colistin, anti-respiratory syncytial virus, lymphocyte immune globulin (Thymoglovin, Atgam), panorex, alpha-antitrypsin, botulinin, lung surfactant protein, tumor necrosis receptor-IgG fusion protein (enbrel), gonadorelin, proteins of influenza vaccine, proteins of rotavirus vaccine, proteins of haemophilus b conjugate vaccine, proteins of poliovirus vaccine, proteins of pneumococcal conjugate vaccine, proteins of meningococcal C vaccine, proteins of influenza vaccine, megakaryocyte growth and development factor (MGDF), neuroimmunophilin ligand-A (NIL-A), brain-derived neurotrophic factor (BDNF), glial cell line-derived neurotrophic factor (GDNF), leptin (native), leptin B, leptin C, IL-1RA (interleukin-1RA), R-568, novel erythropoiesis-stimulating protein (NESP), humanized mAb to rous sarcoma virus (MEDI-493), glutamyl-tryptophan dipeptide IM862, LFA-3TIP immunosuppressive, 5 humanized anti-CD40-ligand monoclonal antibody (5c8), gelsonin enzyme, tissue factor pathway inhibitor (TFPI), proteins of meningitis B vaccine, antimetastatic cancer antibody (mAb 17-1A), chimeric (human & mouse) mAb against TNF α , mAb against factor VII, relaxin, capreomycin, glycopeptide (LY333328), recombinant human activated protein C (rhAPC), humanized mAb against the epidermal growth 10 factor receptor-2, alteplase, anti-CD20 antigen, C2B8 antibody, insulin-like growth factor-1, atrial natriuretic peptide (anaritide), tenectaplasme, anti-CD11a antibody (hu 1124), anti-CD18 antibody, mAb LDP-02, anti-VEGF antibody, fab fragment of anti-VEGF Ab, APO2 ligand (tumor necrosis factor-related apoptosis-inducing ligand), rTGF- β (transforming growth factor- β), alpha-antitrypsin, ananain (a pineapple enzyme), 15 humanized mAb CTLA4IG, PRO 542 (mAb), D2E7 (mAb), calf intestine alkaline phosphatase, α -L-iduronidase, α -L-galactosidase (humanglutamic acid decarboxylase, acid sphingomyelinase, bone morphogenetic protein-2 (rhBMP-2), proteins of HIV vaccine, T cell receptor (TCR) peptide vaccine, TCR peptides, V beta 3 and V beta 13.1. (IR502), (IR501), BI 1050/1272 mAb against very late antigen-4 (VLA-4), 20 C225 humanized mAb to EGF receptor, anti-idiotype antibody to GD3 glycolipid, antibacterial peptide against *H. pylori*, MDX-447 bispecific humanized mAb to EGF receptor, anti-cytomegalovirus (CMV), Medi-491 B19 parvovirus vaccine, humanized recombinant mAb (IgG1k) against respiratory syncytial virus (RSV), urinary tract 25 30

infection vaccine (against "pili" on *Escherechia coli* strains), proteins of lyme disease vaccine against *B. burgdorferi* protein (DbpA), proteins of Medi-501 human papilloma virus-11 vaccine (HPV), *Streptococcus pneumoniae* vaccine, Medi-507 mAb (humanized form of BTI-322) against CD2 receptor on T-cells, MDX-33 mAb to Fc_YR1 receptor, MDX-RA immunotoxin (ricin A linked) mAb, MDX-210 bi-specific mAb against HER-2, MDX-447 bi-specific mAb against EGF receptor, MDX-22, MDX-220 bi-specific mAb against TAG-72 on tumors, colony-stimulating factor (CSF) (molgramostim), humanized mAb to the IL-2 R α -chain (basiliximab), mAb to IgE (IGE 025A), myelin basic protein-altered peptide (MSP771A), 10 humanized mAb against the epidermal growth receptor-2, humanized mAb against the α subunit of the interleukin-2 receptor, low molecular weight heparin, anti-hemophilic factor, and bactericidal/permeability-increasing protein (r-BPI).

The peptides and proteins made using the present invention may be labeled using labels and techniques known to one of ordinary skill in the art. Some of these 15 labels are described in the "Handbook of Fluorescent Probes and Research Products", ninth edition, Richard P. Haugland (ed) Molecular Probes, Inc. Eugene, OR), which is incorporated herein in its entirety. Some of these labels may be genetically engineered into the polynucleotide sequence for the expression of the selected protein or peptide. The peptides and proteins may also have label-incorporation "handles" 20 incorporated to allow labeling of an otherwise difficult or impossible to label protein.

It is to be understood that the various classes of desired peptides and proteins, as well as specific peptides and proteins described in this section may be modified as described below by inserting selected codons for desired amino acid substitutions into the gene incorporated into the transgenic animal.

25 The present invention may also be used to produce desired molecules other than proteins and peptides including, but not limited to, lipoproteins such as high density lipoprotein (HDL), HDL-Milano, and low density lipoprotein, lipids, carbohydrates, siRNA and ribozymes. In these embodiments, a gene of interest encodes a nucleic acid molecule or a protein that directs production of the desired 30 molecule.

The present invention further encompasses the use of inhibitory molecules to inhibit endogenous (i.e., non-vector) protein production. These inhibitory molecules include antisense nucleic acids, siRNA and inhibitory proteins. In one embodiment, a

transposon-based vector containing an ovalbumin DNA sequence, that upon transcription forms a double stranded RNA molecule, is transfected into an animal such as a bird and the bird's production of endogenous ovalbumin protein is reduced by the interference RNA mechanism (RNAi). Additionally, inducible knockouts or 5 knockdowns of the endogenous protein may be created to achieve a reduction or inhibition of endogenous protein production.

Modified Desired Proteins and Peptides

"Proteins", "peptides," "polypeptides" and "oligopeptides" are chains of amino acids (typically L-amino acids) whose alpha carbons are linked through peptide bonds 10 formed by a condensation reaction between the carboxyl group of the alpha carbon of one amino acid and the amino group of the alpha carbon of another amino acid. The terminal amino acid at one end of the chain (i.e., the amino terminal) has a free amino group, while the terminal amino acid at the other end of the chain (i.e., the carboxy terminal) has a free carboxyl group. As such, the term "amino terminus" (abbreviated 15 N-terminus) refers to the free alpha-amino group on the amino acid at the amino terminal of the protein, or to the alpha-amino group (imino group when participating in a peptide bond) of an amino acid at any other location within the protein. Similarly, the term "carboxy terminus" (abbreviated C-terminus) refers to the free 20 carboxyl group on the amino acid at the carboxy terminus of a protein, or to the carboxyl group of an amino acid at any other location within the protein.

Typically, the amino acids making up a protein are numbered in order, starting at the amino terminal and increasing in the direction toward the carboxy terminal of the protein. Thus, when one amino acid is said to "follow" another, that amino acid is positioned closer to the carboxy terminal of the protein than the preceding amino acid. 25 The term "residue" is used herein to refer to an amino acid (D or L) or an amino acid mimetic that is incorporated into a protein by an amide bond. As such, the amino acid may be a naturally occurring amino acid or, unless otherwise limited, may encompass known analogs of natural amino acids that function in a manner similar to the naturally occurring amino acids (i.e., amino acid mimetics). Moreover, an amide 30 bond mimetic includes peptide backbone modifications well known to those skilled in the art.

Furthermore, one of skill will recognize that, as mentioned above, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a

small percentage of amino acids (typically less than about 5%, more typically less than about 1%) in an encoded sequence are conservatively modified variations where the alterations result in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. The following six groups each contain amino acids that are conservative substitutions for one another:

- 5 1) Alanine (A), Serine (S), Threonine (T);
- 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 10 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

A conservative substitution is a substitution in which the substituting amino acid (naturally occurring or modified) is structurally related to the amino acid being substituted, i.e., has about the same size and electronic properties as the amino acid being substituted. Thus, the substituting amino acid would have the same or a similar functional group in the side chain as the original amino acid. A "conservative substitution" also refers to utilizing a substituting amino acid which is identical to the amino acid being substituted except that a functional group in the side chain is 20 protected with a suitable protecting group.

Suitable protecting groups are described in Green and Wuts, "Protecting Groups in Organic Synthesis", John Wiley and Sons, Chapters 5 and 7, 1991, the teachings of which are incorporated herein by reference. Preferred protecting groups are those which facilitate transport of the peptide through membranes, for example, by 25 reducing the hydrophilicity and increasing the lipophilicity of the peptide, and which can be cleaved, either by hydrolysis or enzymatically (Ditter et al., 1968. J. Pharm. Sci. 57:783; Ditter et al., 1968. J. Pharm. Sci. 57:828; Ditter et al., 1969. J. Pharm. Sci. 58:557; King et al., 1987. Biochemistry 26:2294; Lindberg et al., 1989. Drug Metabolism and Disposition 17:311; Tunek et al., 1988. Biochem. Pharm. 37:3867; 30 Anderson et al., 1985 Arch. Biochem. Biophys. 239:538; and Singhal et al., 1987. FASEB J. 1:220). Suitable hydroxyl protecting groups include ester, carbonate and carbamate protecting groups. Suitable amine protecting groups include acyl groups and alkoxy or aryloxy carbonyl groups, as described above for N-terminal protecting

groups. Suitable carboxylic acid protecting groups include aliphatic, benzyl and aryl esters, as described below for C-terminal protecting groups. In one embodiment, the carboxylic acid group in the side chain of one or more glutamic acid or aspartic acid residues in a peptide of the present invention is protected, preferably as a methyl, 5 ethyl, benzyl or substituted benzyl ester, more preferably as a benzyl ester.

Provided below are groups of naturally occurring and modified amino acids in which each amino acid in a group has similar electronic and steric properties. Thus, a conservative substitution can be made by substituting an amino acid with another amino acid from the same group. It is to be understood that these groups are non-10 limiting, i.e. that there are additional modified amino acids which could be included in each group.

Group I includes leucine, isoleucine, valine, methionine and modified amino acids having the following side chains: ethyl, n-propyl n-butyl. Preferably, Group I includes leucine, isoleucine, valine and methionine.

15 Group II includes glycine, alanine, valine and a modified amino acid having an ethyl side chain. Preferably, Group II includes glycine and alanine.

Group III includes phenylalanine, phenylglycine, tyrosine, tryptophan, cyclohexylmethyl glycine, and modified amino residues having substituted benzyl or phenyl side chains. Preferred substituents include one or more of 20 the following: halogen, methyl, ethyl, nitro, —NH₂, methoxy, ethoxy and —CN. Preferably, Group III includes phenylalanine, tyrosine and tryptophan.

Group IV includes glutamic acid, aspartic acid, a substituted or unsubstituted aliphatic, aromatic or benzylic ester of glutamic or aspartic acid (e.g., methyl, 25 ethyl, n-propyl iso-propyl, cyclohexyl, benzyl or substituted benzyl), glutamine, asparagine, —CO—NH— alkylated glutamine or asparagines (e.g., methyl, ethyl, n-propyl and iso-propyl) and modified amino acids having the side chain —(CH₂)₃—COOH, an ester thereof (substituted or unsubstituted aliphatic, aromatic or benzylic ester), an amide thereof and a substituted or 30 unsubstituted N-alkylated amide thereof. Preferably, Group IV includes glutamic acid, aspartic acid, methyl aspartate, ethyl aspartate, benzyl aspartate and methyl glutamate, ethyl glutamate and benzyl glutamate, glutamine and asparagine.

Group V includes histidine, lysine, ornithine, arginine, N-nitroarginine, β -cycloarginine, γ -hydroxyarginine, N-amidinocitruline and 2-amino-4-guanidinobutanoic acid, homologs of lysine, homologs of arginine and homologs of ornithine. Preferably, Group V includes histidine, lysine, arginine and ornithine. A homolog of an amino acid includes from 1 to about 3 additional or subtracted methylene units in the side chain.

5 Group VI includes serine, threonine, cysteine and modified amino acids having C1-C5 straight or branched alkyl side chains substituted with —OH or —SH, for example, —CH₂CH₂OH, —CH₂CH₂CH₂OH or -CH₂CH₂OHCH₃. Preferably, 10 Group VI includes serine, cysteine or threonine.

In another aspect, suitable substitutions for amino acid residues include "severe" substitutions. A "severe substitution" is a substitution in which the substituting amino acid (naturally occurring or modified) has significantly different size and/or electronic properties compared with the amino acid being substituted.

15 Thus, the side chain of the substituting amino acid can be significantly larger (or smaller) than the side chain of the amino acid being substituted and/or can have functional groups with significantly different electronic properties than the amino acid being substituted. Examples of severe substitutions of this type include the substitution of phenylalanine or cyclohexylmethyl glycine for alanine, isoleucine for 20 glycine, a D amino acid for the corresponding L amino acid, or —NH—CH[(-CH₂)₅—COOH]—CO— for aspartic acid. Alternatively, a functional group may be added to the side chain, deleted from the side chain or exchanged with another functional group. Examples of severe substitutions of this type include adding of valine, leucine or isoleucine, exchanging the carboxylic acid in the side chain of 25 aspartic acid or glutamic acid with an amine, or deleting the amine group in the side chain of lysine or ornithine. In yet another alternative, the side chain of the substituting amino acid can have significantly different steric and electronic properties that the functional group of the amino acid being substituted. Examples of such modifications include tryptophan for glycine, lysine for aspartic acid and — 30 (CH₂)₄COOH for the side chain of serine. These examples are not meant to be limiting.

In another embodiment, for example in the synthesis of a peptide 26 amino acids in length, the individual amino acids may be substituted according in the following manner:

AA₁ is serine, glycine, alanine, cysteine or threonine;

5 AA₂ is alanine, threonine, glycine, cysteine or serine;

AA₃ is valine, arginine, leucine, isoleucine, methionine, ornithine, lysine, N-nitroarginine, β -cycloarginine, γ -hydroxyarginine, N-amidinocitruline or 2-amino-4-guanidinobutanoic acid;

AA₄ is proline, leucine, valine, isoleucine or methionine;

10 AA₅ is tryptophan, alanine, phenylalanine, tyrosine or glycine;

AA₆ is serine, glycine, alanine, cysteine or threonine;

AA₇ is proline, leucine, valine, isoleucine or methionine;

AA₈ is alanine, threonine, glycine, cysteine or serine;

AA₉ is alanine, threonine, glycine, cysteine or serine;

15 AA₁₀ is leucine, isoleucine, methionine or valine;

AA₁₁ is serine, glycine, alanine, cysteine or threonine;

AA₁₂ is leucine, isoleucine, methionine or valine;

AA₁₃ is leucine, isoleucine, methionine or valine;

AA₁₄ is glutamine, glutamic acid, aspartic acid, asparagine, or a substituted or 20 unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;

AA₁₅ is arginine, N-nitroarginine, β -cycloarginine, γ -hydroxy-arginine, N-amidinocitruline or 2-amino-4-guanidino-butanoic acid

AA₁₆ is proline, leucine, valine, isoleucine or methionine;

AA₁₇ is serine, glycine, alanine, cysteine or threonine;

25 AA₁₈ is glutamic acid, aspartic acid, asparagine, glutamine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;

AA₁₉ is aspartic acid, asparagine, glutamic acid, glutamine, leucine, valine, isoleucine, methionine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;

30 AA₂₀ is valine, arginine, leucine, isoleucine, methionine, ornithine, lysine, N-nitroarginine, β -cycloarginine, γ -hydroxyarginine, N-amidinocitruline or 2-amino-4-guanidinobutanoic acid;

AA₂₁ is alanine, threonine, glycine, cysteine or serine;

AA₂₂ is alanine, threonine, glycine, cysteine or serine;
AA₂₃ is histidine, serine, threonine, cysteine, lysine or ornithine;
AA₂₄ is threonine, aspartic acid, serine, glutamic acid or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;
5 AA₂₅ is asparagine, aspartic acid, glutamic acid, glutamine, leucine, valine, isoleucine, methionine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid; and
AA₂₆ is cysteine, histidine, serine, threonine, lysine or ornithine.

It is to be understood that these amino acid substitutions may be made for
10 longer or shorter peptides than the 26 mer in the preceding example above, and for proteins.

In one embodiment of the present invention, codons for the first several N-terminal amino acids of the transposase are modified such that the third base of each codon is changed to an A or a T without changing the corresponding amino acid. It is
15 preferable that between approximately 1 and 20, more preferably 3 and 15, and most preferably between 4 and 12 of the first N-terminal codons of the gene of interest are modified such that the third base of each codon is changed to an A or a T without changing the corresponding amino acid. In one embodiment, the first ten N-terminal codons of the gene of interest are modified in this manner.

20 When several desired proteins, protein fragments or peptides are encoded in the gene of interest to be incorporated into the genome, one of skill in the art will appreciate that the proteins, protein fragments or peptides may be separated by a spacer molecule such as, for example, a peptide, consisting of one or more amino acids. Generally, the spacer will have no specific biological activity other than to join
25 the desired proteins, protein fragments or peptides together, or to preserve some minimum distance or other spatial relationship between them. However, the constituent amino acids of the spacer may be selected to influence some property of the molecule such as the folding, net charge, or hydrophobicity. The spacer may also be contained within a nucleotide sequence with a purification handle or be flanked by
30 proteolytic cleavage sites.

Such polypeptide spacers may have from about 5 to about 40 amino acid residues. The spacers in a polypeptide are independently chosen, but are preferably all the same. The spacers should allow for flexibility of movement in space and are

therefore typically rich in small amino acids, for example, glycine, serine, proline or alanine. Preferably, peptide spacers contain at least 60%, more preferably at least 80% glycine or alanine. In addition, peptide spacers generally have little or no biological and antigenic activity. Preferred spacers are (Gly-Pro-Gly-Gly)_x (SEQ ID NO:5) and (Gly₄-Ser)_y, wherein x is an integer from about 3 to about 9 and y is an integer from about 1 to about 8. Specific examples of suitable spacers include (Gly-Pro-Gly-Gly)₃

SEQ ID NO:6 Gly Pro Gly Gly Gly Pro Gly Gly Pro Gly Gly
(Gly₄-Ser)₃

10 SEQ ID NO:7 Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
or (Gly₄-Ser)₄

SEQ ID NO:8 Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
Gly Gly Gly Gly Ser.

Nucleotide sequences encoding for the production of residues which may be useful in purification of the expressed recombinant protein may also be built into the vector. Such sequences are known in the art and include the glutathione binding domain from glutathione S-transferase, polylysine, hexa-histidine or other cationic amino acids, thioredoxin, hemagglutinin antigen and maltose binding protein.

Additionally, nucleotide sequences may be inserted into the gene of interest to be incorporated so that the protein or peptide can also include from one to about six amino acids that create signals for proteolytic cleavage. In this manner, if a gene is designed to make one or more peptides or proteins of interest in the transgenic animal, specific nucleotide sequences encoding for amino acids recognized by enzymes may be incorporated into the gene to facilitate cleavage of the large protein or peptide sequence into desired peptides or proteins or both. For example, nucleotides encoding a proteolytic cleavage site can be introduced into the gene of interest so that a signal sequence can be cleaved from a protein or peptide encoded by the gene of interest. Nucleotide sequences encoding other amino acid sequences which display pH sensitivity or chemical sensitivity may also be added to the vector to facilitate separation of the signal sequence from the peptide or protein of interest.

In one embodiment of the present invention, a TAG sequence is linked to the gene of interest. The TAG sequence serves three purposes: 1) it allows free rotation of the peptide or protein to be isolated so there is no interference from the native

protein or signal sequence, i.e. vitellogenin, 2) it provides a "purification handle" to isolate the protein using column purification, and 3) it includes a cleavage site to remove the desired protein from the signal and purification sequences. Accordingly, as used herein, a TAG sequence includes a spacer sequence, a purification handle and 5 a cleavage site. The spacer sequences in the TAG proteins contain one or more repeats shown in SEQ ID NO:25. A preferred spacer sequence comprises the sequence provided in SEQ ID NO:26. One example of a purification handle is the gp41 hairpin loop from HIV I. Exemplary gp41 polynucleotide and polypeptide sequences are provided in SEQ ID NO:24 and SEQ ID NO:23, respectively. 10 However, it should be understood that any antigenic region may be used as a purification handle, including any antigenic region of gp41. Preferred purification handles are those that elicit highly specific antibodies. Additionally, the cleavage site can be any protein cleavage site known to one of ordinary skill in the art and includes an enterokinase cleavage site comprising the Asp Asp Asp Asp Lys sequence (SEQ 15 ID NO:9) and a furin cleavage site. Constructs containing a TAG sequence are shown in Figures 2 and 3. In one embodiment of the present invention, the TAG sequence comprises a polynucleotide sequence of SEQ ID NO:22.

Methods of Administering Transposon-Based Vectors

In addition to the transposon-based vectors described above, the present 20 invention also includes methods of administering the transposon-based vectors to an animal, methods of producing a transgenic animal wherein a gene of interest is incorporated into the germline of the animal and methods of producing a transgenic animal wherein a gene of interest is incorporated into cells other than the germline cells of the animal. The transposon-based vectors of the present invention may be 25 administered to an animal via any method known to those of skill in the art, including, but not limited to, intraembryonic, intratesticular, intraoviduct, intraperitoneal, intraarterial, intravenous, topical, oral, nasal, and pronuclear injection methods of administration, or any combination thereof. The transposon-based vectors may also be administered within the lumen of an organ, into an organ, into a body cavity, into 30 the cerebrospinal fluid, through the urinary system or through any route to reach the desired cells.

The transposon-based vectors may be delivered through the vascular system to be distributed to the cells supplied by that vessel. For example, the compositions may

be placed in the artery supplying the ovary or supplying the fallopian tube to transfect cells in those tissues. In this manner, follicles could be transfected to create a germline transgenic animal. Alternatively, supplying the compositions through the artery leading to the oviduct would preferably transfect the tubular gland and

5 epithelial cells. Such transfected cells could manufacture a desired protein or peptide for deposition in the egg white. Administration of the compositions through the portal vein would target uptake and transformation of hepatic cells. Administration through the urethra and into the bladder would target the transitional epithelium of the bladder. Administration through the vagina and cervix would target the lining of the uterus.

10 Administration through the internal mammary artery would transfect secretory cells of the lactating mammary gland to perform a desired function, such as to synthesize and secrete a desired protein or peptide into the milk.

In a preferred embodiment, the animal is an egg-laying animal, and more preferably, an avian. In one embodiment, between approximately 1 and 50 µg, 15 preferably between 1 and 20 µg, and more preferably between 5 and 10 µg of transposon-based vector DNA is administered to the oviduct of a bird. Optimal ranges depending upon the type of bird and the bird's stage of sexual maturity. Intraoviduct administration of the transposon-based vectors of the present invention result in a PCR positive signal in the oviduct tissue, whereas intravascular 20 administration results in a PCR positive signal in the liver. In other embodiments, the transposon-based vector is administered to an artery that supplies the oviduct or the liver. These methods of administration may also be combined with any methods for facilitating transfection, including without limitation, electroporation, gene guns, injection of naked DNA, and use of dimethyl sulfoxide (DMSO).

25 The present invention includes a method of intraembryonic administration of a transposon-based vector to an avian embryo comprising the following steps: 1) incubating an egg on its side at room temperature for two hours to allow the embryo contained therein to move to top dead center (TDC); 2) drilling a hole through the shell without penetrating the underlying shell membrane; 3) injecting the embryo with 30 the transposon-based vector in solution; 4) sealing the hole in the egg; and 5) placing the egg in an incubator for hatching. Administration of the transposon-based vector can occur anytime between immediately after egg lay (when the embryo is at Stage X) and hatching. Preferably, the transposon-based vector is administered between 1 and

7 days after egg lay, more preferably between 1 and 2 days after egg lay. The transposon-based vectors may be introduced into the embryo in amounts ranging from about 5.0 μ g to 10 μ g, preferably 1.0 μ g to 100 μ g. Additionally, the transposon-based vector solution volume may be between approximately 1 μ l to 75 μ l in quail
5 and between approximately 1 μ l to 500 μ l in chicken.

The present invention also includes a method of intratesticular administration of a transposon-based vector including injecting a bird with a composition comprising the transposon-based vector, an appropriate carrier and an appropriate transfection reagent. In one embodiment, the bird is injected before sexual maturity, preferably
10 between approximately 4-14 weeks, more preferably between approximately 6-14 weeks and most preferably between 8-12 weeks old. In another embodiment, a mature bird is injected with a transposon-based vector an appropriate carrier and an appropriate transfection reagent. The mature bird may be any type of bird, but in one example the mature bird is a quail.

15 A bird is preferably injected prior to the development of the blood-testis barrier, which thereby facilitates entry of the transposon-based vector into the seminiferous tubules and transfection of the spermatogonia or other germline cells. At and between the ages of 4, 6, 8, 10, 12, and 14 weeks, it is believed that the testes of chickens are likely to be most receptive to transfection. In this age range, the
20 blood/testis barrier has not yet formed, and there is a relatively high number of spermatogonia relative to the numbers of other cell types, e.g., spermatids, etc. See J. Kumaran et al., 1949. Poultry Sci., 29:511-520. See also E. Oakberg, 1956. Am. J. Anatomy, 99:507-515; and P. Kluin et al., 1984. Anat. Embryol., 169:73-78.

The transposon-based vectors may be introduced into a testis in an amount
25 ranging from about 0.1 μ g to 10 μ g, preferably 1 μ g to 10 μ g, more preferably 3 μ g to 10 μ g. In a quail, about 5 μ g is a preferred amount. In a chicken, about 5 μ g to 10 μ g per testis is preferred. These amounts of vector DNA may be injected in one dose or multiple doses and at one site or multiple sites in the testis. In a preferred embodiment, the vector DNA is administered at multiple sites in a single testis, both
30 testes being injected in this manner. In one embodiment, injection is spread over three injection sites: one at each end of the testis, and one in the middle. Additionally, the transposon-based vector solution volume may be between approximately 1 μ l to 75 μ l in quail and between approximately 1 μ l to 500 μ l in chicken. In a preferred

embodiment, the transposon-based vector solution volume may be between approximately 20 μ l to 60 μ l in quail and between approximately 50 μ l to 250 μ l in chicken. Both the amount of vector DNA and the total volume injected into each testis may be determined based upon the age and size of the bird.

5 According to the present invention, the transposon-based vector is administered in conjunction with an acceptable carrier and/or transfection reagent. Acceptable carriers include, but are not limited to, water, saline, Hanks Balanced Salt Solution (HBSS), Tris-EDTA (TE) and lyotropic liquid crystals. Transfection reagents commonly known to one of ordinary skill in the art that may be employed
10 include, but are not limited to, the following: cationic lipid transfection reagents, cationic lipid mixtures, polyamine reagents, liposomes and combinations thereof; SUPERFECT \circledR , Cytofectene, BioPORTER \circledR , GenePORTER \circledR , NeuroPORTER \circledR , and perfectin from Gene Therapy Systems; lipofectamine, cellfectin, DMRIE-C oligofectamine, and PLUS reagent from InVitrogen; Xtreme gene, fugene, DOSPER
15 and DOTAP from Roche; Lipotaxi and Genejammer from Strategene; and Escort from SIGMA. In one embodiment, the transfection reagent is SUPERFECT \circledR . The ratio of DNA to transfection reagent may vary based upon the method of administration. In one embodiment, the transposon-based vector is administered intratesticularly and the ratio of DNA to transfection reagent can be from 1:1.5 to
20 1:15, preferably 1:2 to 1:10, all expressed as wt/vol. Transfection may also be accomplished using other means known to one of ordinary skill in the art, including without limitation electroporation, gene guns, injection of naked DNA, and use of dimethyl sulfoxide (DMSO).

Depending upon the cell or tissue type targeted for transfection, the form of
25 the transposon-based vector may be important. Plasmids harvested from bacteria are generally closed circular supercoiled molecules, and this is the preferred state of a vector for gene delivery because of the ease of preparation. In some instances, transposase expression and insertion may be more efficient in a relaxed, closed circular configuration or in a linear configuration. In still other instances, a purified
30 transposase protein may be co-injected with a transposon-based vector containing the gene of interest for more immediate insertion. This could be accomplished by using a transfection reagent complexed with both the purified transposase protein and the transposon-based vector.

Testing for and Breeding Animals Carrying the Transgene

Following administration of a transposon-based vector to an animal, DNA is extracted from the animal to confirm integration of the gene of interest. Actual frequencies of integration are estimated both by comparative strength of the PCR signal, and by histological evaluation of the tissues by quantitative PCR. Another method for estimating the rate of transgene insertion is the so-called primed in situ hybridization technique (PRINS). This method determines not only which cells carry a transgene of interest, but also into which chromosome the gene has inserted, and even what portion of the chromosome. Briefly, labeled primers are annealed to chromosome spreads (affixed to glass slides) through one round of PCR, and the slides are then developed through normal in situ hybridization procedures. This technique combines the best features of in situ PCR and fluorescence in situ hybridization (FISH) to provide distinct chromosome location and copy number of the gene in question. The 28s rRNA gene will be used as a positive control for spermatogonia to confirm that the technique is functioning properly. Using different fluorescent labels for the transgene and the 28s gene causes cells containing a transgene to fluoresce with two different colored tags.

Breeding experiments are also conducted to determine if germline transmission of the transgene has occurred. In a general bird breeding experiment performed according to the present invention, each male bird was exposed to 2-3 different adult female birds for 3-4 days each. This procedure was continued with different females for a total period of 6-12 weeks. Eggs were collected daily for up to 14 days after the last exposure to the transgenic male, and each egg was incubated in a standard incubator. In the first series of experiments the resulting embryos were examined for transgene presence at day 3 or 4 using PCR.

Any male producing a transgenic embryo was bred to additional females. Eggs from these females were incubated, hatched, and the chicks tested for the exogenous DNA. Any embryos that died were necropsied and examined directly for the transgene or protein encoded by the transgene, either by fluorescence or PCR. The offspring that hatched and were found to be positive for the exogenous DNA were raised to maturity. These birds were bred to produce further generations of transgenic birds, to verify efficiency of the transgenic procedure and the stable

incorporation of the transgene into the germ line. The resulting embryos were examined for transgene presence at day 3 or 4 using PCR.

It is to be understood that the above procedure can be modified to suit animals other than birds and that selective breeding techniques may be performed to amplify

5 gene copy numbers and protein output.

Production of Desired Proteins or Peptides in Egg White

In one embodiment, the transposon-based vectors of the present invention may be administered to a bird for production of desired proteins or peptides in the egg white. These transposon-based vectors preferably contain one or more of an

10 ovalbumin promoter, an ovomucoid promoter, an ovalbumin signal sequence and an ovomucoid signal sequence. Oviduct-specific ovalbumin promoters are described in B. O'Malley et al., 1987. EMBO J., vol. 6, pp. 2305-12; A. Qiu et al., 1994. Proc. Nat. Acad. Sci. (USA), vol. 91, pp. 4451-4455; D. Monroe et al., 2000. Biochim. Biophys. Acta, 1517 (1):27-32; H. Park et al., 2000. Biochem., 39:8537-8545; and T.

15 Muramatsu et al., 1996. Poult. Avian Biol. Rev., 6:107-123. Examples of transposon-based vectors designed for production of a desired protein in an egg white are shown in Figures 2 and 3.

Production of Desired Proteins or Peptides in Egg Yolk

The present invention is particularly advantageous for production of recombinant peptides and proteins of low solubility in the egg yolk. Such proteins include, but are not limited to, membrane-associated or membrane-bound proteins, lipophilic compounds; attachment factors, receptors, and components of second messenger transduction machinery. Low solubility peptides and proteins are particularly challenging to produce using conventional recombinant protein

25 production techniques (cell and tissue cultures) because they aggregate in water-based, hydrophilic environments. Such aggregation necessitates denaturation and refolding of the recombinantly-produced proteins, which may deleteriously affect their structure and function. Moreover, even highly soluble recombinant peptides and proteins may precipitate and require denaturation and renaturation when produced in 30 sufficiently high amounts in recombinant protein production systems. The present invention provides an advantageous resolution of the problem of protein and peptide solubility during production of large amounts of recombinant proteins.

In one embodiment of the present invention, deposition of a desired protein into the egg yolk is accomplished by attaching a sequence encoding a protein capable of binding to the yolk vitellogenin receptor to a gene of interest that encodes a desired protein. This transposon-based vector can be used for the receptor-mediated uptake 5 of the desired protein by the oocytes. In a preferred embodiment, the sequence ensuring the binding to the vitellogenin receptor is a targeting sequence of a vitellogenin protein. The invention encompasses various vitellogenin proteins and their targeting sequences. In a preferred embodiment, a chicken vitellogenin protein targeting sequence is used, however, due to the high degree of conservation among 10 vitellogenin protein sequences and known cross-species reactivity of vitellogenin targeting sequences with their egg-yolk receptors, other vitellogenin targeting sequences can be substituted. One example of a construct for use in the transposon-based vectors of the present invention and for deposition of an insulin protein in an egg yolk is provided in SEQ ID NO:27. In this embodiment, the transposon-based 15 vector contains a vitellogenin promoter, a vitellogenin targeting sequence, a TAG sequence, a pro-insulin sequence and a synthetic polyA sequence. The present invention includes, but is not limited to, vitellogenin targeting sequences residing in the N-terminal domain of vitellogenin, particularly in lipovitellin I. In one embodiment, the vitellogenin targeting sequence contains the polynucleotide 20 sequence of SEQ ID NO:18.

In a preferred embodiment, the transposon-based vector contains a transposase gene operably-linked to a liver-specific promoter and a gene of interest operably-linked to a liver-specific promoter and a vitellogenin targeting sequence. Figure 4 shows an example of such a construct. In another preferred embodiment, the 25 transposon-based vector contains a transposase gene operably-linked to a constitutive promoter and a gene of interest operably-linked to a liver-specific promoter and a vitellogenin targeting sequence.

Isolation and Purification of Desired Protein or Peptide

For large-scale production of protein, an animal breeding stock that is 30 homozygous for the transgene is preferred. Such homozygous individuals are obtained and identified through, for example, standard animal breeding procedures or PCR protocols.

Once expressed, peptides, polypeptides and proteins can be purified according to standard procedures known to one of ordinary skill in the art, including ammonium sulfate precipitation, affinity columns, column chromatography, gel electrophoresis, high performance liquid chromatography, immunoprecipitation and the like.

5 Substantially pure compositions of about 50 to 99% homogeneity are preferred, and 80 to 95% or greater homogeneity are most preferred for use as therapeutic agents.

In one embodiment of the present invention, the animal in which the desired protein is produced is an egg-laying animal. In a preferred embodiment of the present invention, the animal is an avian and a desired peptide, polypeptide or protein is 10 isolated from an egg white. Egg white containing the exogenous protein or peptide is separated from the yolk and other egg constituents on an industrial scale by any of a variety of methods known in the egg industry. See, e.g., W. Stadelman et al. (Eds.), Egg Science & Technology, Haworth Press, Binghamton, NY (1995). Isolation of the exogenous peptide or protein from the other egg white constituents is accomplished 15 by any of a number of polypeptide isolation and purification methods well known to one of ordinary skill in the art. These techniques include, for example, chromatographic methods such as gel permeation, ion exchange, affinity separation, metal chelation, HPLC, and the like, either alone or in combination. Another means that may be used for isolation or purification, either in lieu of or in addition to 20 chromatographic separation methods, includes electrophoresis. Successful isolation and purification is confirmed by standard analytic techniques, including HPLC, mass spectroscopy, and spectrophotometry. These separation methods are often facilitated if the first step in the separation is the removal of the endogenous ovalbumin fraction of egg white, as doing so will reduce the total protein content to be further purified by 25 about 50%.

To facilitate or enable purification of a desired protein or peptide, transposon-based vectors may include one or more additional epitopes or domains. Such epitopes or domains include DNA sequences encoding enzymatic or chemical cleavage sites including, but not limited to, an enterokinase cleavage site; the glutathione binding 30 domain from glutathione S-transferase; polylysine; hexa-histidine or other cationic amino acids; thioredoxin; hemagglutinin antigen; maltose binding protein; a fragment of gp41 from HIV; and other purification epitopes or domains commonly known to one of skill in the art.

In one representative embodiment, purification of desired proteins from egg white utilizes the antigenicity of the ovalbumin carrier protein and particular attributes of a TAG linker sequence that spans ovalbumin and the desired protein. The TAG sequence is particularly useful in this process because it contains 1) a highly antigenic 5 epitope, a fragment of gp41 from HIV, allowing for stringent affinity purification, and, 2) a recognition site for the protease enterokinase immediately juxtaposed to the desired protein. In a preferred embodiment, the TAG sequence comprises approximately 50 amino acids. A representative TAG sequence is provided below.

10 Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Thr Thr Cys Ile Leu Lys Gly Ser Cys
Gly Trp Ile Gly Leu Leu Asp Asp Asp Asp Lys (SEQ ID NO:22)

The underlined sequences were taken from the hairpin loop domain of HIV gp-41 15 (SEQ ID NO:23). Sequences in italics represent the cleavage site for enterokinase (SEQ ID NO:9). The spacer sequence upstream of the loop domain was made from repeats of (Pro Ala Asp Asp Ala) (SEQ ID NO:25) to provide free rotation and promote surface availability of the hairpin loop from the ovalbumin carrier protein.

Isolation and purification of a desired protein is performed as follows:

- 20 1. Enrichment of the egg white protein fraction containing ovalbumin and the transgenic ovalbumin-TAG-desired protein.
2. Size exclusion chromatography to isolate only those proteins within a narrow range of molecular weights (a further enrichment of step 1).
3. Ovalbumin affinity chromatography. Highly specific antibodies to ovalbumin 25 will eliminate virtually all extraneous egg white proteins except ovalbumin and the transgenic ovalbumin-TAG-desired protein.
4. gp41 affinity chromatography using anti-gp41 antibodies. Stringent application of this step will result in virtually pure transgenic ovalbumin-TAG-desired protein.
- 30 5. Cleavage of the transgene product can be accomplished in at least one of two ways:
 - a. The transgenic ovalbumin-TAG-desired protein is left attached to the gp41 affinity resin (beads) from step 4 and the protease enterokinase is

added. This liberates the transgene target protein from the gp41 affinity resin while the ovalbumin-TAG sequence is retained. Separation by centrifugation (in a batch process) or flow through (in a column purification), leaves the desired protein together with enterokinase in solution. Enterokinase is recovered and reused.

5

- b. Alternatively, enterokinase is immobilized on resin (beads) by the addition of poly-lysine moieties to a non-catalytic area of the protease. The transgenic ovalbumin-TAG-desired protein eluted from the affinity column of step 4 is then applied to the protease resin. Protease action cleaves the ovalbumin-TAG sequence from the desired protein and leaves both entities in solution. The immobilized enterokinase resin is recharged and reused.
- c. The choice of these alternatives is made depending upon the size and chemical composition of the transgene target protein.

10

and leaves both entities in solution. The immobilized enterokinase resin is recharged and reused.

c. The choice of these alternatives is made depending upon the size and chemical composition of the transgene target protein.

15 6. A final separation of either of these two (5a or 5b) protein mixtures is made using size exclusion, or enterokinase affinity chromatography. This step allows for desalting, buffer exchange and/or polishing, as needed.

Cleavage of the transgene product (ovalbumin-TAG-desired protein) by enterokinase, then, results in two products: ovalbumin-TAG and the desired protein.

20 More specific methods for isolation using the TAG label is provided in the Examples. Some desired proteins may require additions or modifications of the above-described approach as known to one of ordinary skill in the art. The method is scalable from the laboratory bench to pilot and production facility largely because the techniques applied are well documented in each of these settings.

25 It is believed that a typical chicken egg produced by a transgenic animal of the present invention will contain at least 0.001 mg, from about 0.001 to 1.0 mg, or from about 0.001 to 100.0 mg of exogenous protein, peptide or polypeptide, in addition to the normal constituents of egg white (or possibly replacing a small fraction of the latter).

30 One of skill in the art will recognize that after biological expression or purification, the desired proteins, fragments thereof and peptides may possess a conformation substantially different than the native conformations of the proteins, fragments thereof and peptides. In this case, it is often necessary to denature and

reduce protein and then to cause the protein to re-fold into the preferred conformation. Methods of reducing and denaturing proteins and inducing re-folding are well known to those of skill in the art.

Production of Protein or Peptide in Milk

5 In addition to methods of producing eggs containing transgenic proteins or peptides, the present invention encompasses methods for the production of milk containing transgenic proteins or peptides. These methods include the administration of a transposon-based vector described above to a mammal. In one embodiment, the transposon-based vector contains a transposase operably-linked to a constitutive promoter and a gene of interest operably-linked to mammary specific promoter. Genes of interest can include, but are not limited to antiviral and antibacterial proteins 10 and immunoglobulins.

Treatment of Disease and Animal Improvement

In addition to production and isolation of desired molecules, the transposon-based vectors of the present invention can be used for the treatment of various genetic disorders. For example, one or more transposon-based vectors can be administered to a human or animal for the treatment of a single gene disorder including, but not limited to, Huntington's disease, alpha-1-antitrypsin deficiency Alzheimer's disease, various forms of breast cancer, cystic fibrosis, galactosemia, congenital 15 hypothyroidism, maple syrup urine disease, neurofibromatosis 1, phenylketonuria, sickle cell disease, and Smith-Lemli-Opitz (SLO/RSH) Syndrome. Other diseases caused by single gene disorders that may be treated with the present invention include, autoimmune diseases, shipping fever in cattle, mastitis, bacterial or viral diseases, alteration of skin pigment in animals. In these embodiments, the 20 transposon-based vector contains a non-mutated, or non-disease causing form of the gene known to cause such disorder. Preferably, the transposase contained within the transposase-based vector is operably linked to an inducible promoter such as a tissue-specific promoter such that the non-mutated gene of interest is inserted into a specific 25 tissue wherein the mutated gene is expressed in vivo.

30 In one embodiment of the present invention, a transposon-based vector comprising a gene encoding proinsulin is administered to diabetic animals or humans for incorporation into liver cells in order to treat or cure diabetes. The specific incorporation of the proinsulin gene into the liver is accomplished by placing the

transposase gene under the control of liver-specific promoter, such as G6P. This approach is useful for treatment of both Type I and Type II diabetes. The G6P promoter has been shown to be glucose responsive (Arguad, D., et al. 1996. *Diabetes* 45:1563-1571), and thus, glucose-regulated insulin production is achieved using DNA 5 constructs of the present invention. Integrating a proinsulin gene into liver cells circumvents the problem of destruction of pancreatic islet cells in the course of Type I diabetes.

In another embodiment, shortly after diagnosis of Type I diabetes, the cells of the immune system destroying pancreatic β -cells are selectively removed using the 10 transposon-based vectors of the present invention, thus allowing normal β -cells to repopulate the pancreas.

For treatment of Type II diabetes, a transposon-based vector containing a proinsulin gene is specifically incorporated into the pancreas by placing the transposase gene under the control of a pancreas-specific promoter, such as an insulin 15 promoter. In this embodiment, the vector is delivered to a diabetic animal or human via injection into an artery feeding the pancreas. For delivery, the vector is complexed with a transfection agent. The artery distributes the complex throughout the pancreas, where individual cells receive the vector DNA. Following uptake into the target cell, the insulin promoter is recognized by transcriptional machinery of the 20 cell, the transposase encoded by the vector is expressed, and stable integration of the proinsulin gene occurs. It is expected that a small percentage of the transposon-based vector is transported to other tissues, and that these tissues are transfected. However, these tissues are not stably transfected and the proinsulin gene is not incorporated into the cells' DNA due to failure of these cells to activate the insulin promoter. The 25 vector DNA is likely lost when the cell dies or degraded over time.

In other embodiments, one or more transposon-based vectors are administered to an avian for the treatment of a viral or bacterial infection/disease including, but not limited to, Colibacillosis (Coliform infections), Mycoplasmosis (CRD, Air sac, Sinusitis), Fowl Cholera, Necrotic Enteritis, Ulcerative Enteritis (Quail disease), 30 Pullorum Disease, Fowl Typhoid, Botulism, Infectious Coryza, Erysipelas, Avian Pox, Newcastle Disease, Infectious Bronchitis, Quail Bronchitis, Lymphoid Leukosis, Marek's Disease (Visceral Leukosis), Infectious Bursal Disease (Gumboro). In these

embodiments, the transposon-based vectors may be used in a manner similar to traditional vaccines.

In still other embodiments, one or more transposon-based vectors are administered to an animal for the production of an animal with enhanced growth 5 characteristics and nutrient utilization.

The transposon-based vectors of the present invention can be used to transform any animal cell, including but not limited to: cells producing hormones, cytokines, growth factors, or any other biologically active substance; cells of the immune system; cells of the nervous system; muscle (striatal, cardiac, smooth) cells; 10 vascular system cells; endothelial cells; skin cells; mammary cells; and lung cells, including bronchial and alveolar cells. Transformation of any endocrine cell by a transposon-based vector is contemplated as a part of a present invention. In one aspect of the present invention, cells of the immune system may be the target for incorporation of a desired gene or genes encoding for production of antibodies. 15 Accordingly, the thymus, bone marrow, beta lymphocytes (or B cells), gastrointestinal associated lymphatic tissue (GALT), Peyer's patches, bursa Fabricius, lymph nodes, spleen, and tonsil, and any other lymphatic tissue, may all be targets for administration of the compositions of the present invention.

The transposon-based vectors of the present invention can be used to modulate 20 (stimulate or inhibit) production of any substance, including but not limited to a hormone, a cytokine, or a growth factor, by an animal or a human cell. Modulation of a regulated signal within a cell or a tissue, such as production of a second messenger, is also contemplated as a part of the present invention. Use of the transposon-based vectors of the present invention is contemplated for treatment of any animal or human 25 disease or condition that results from underproduction (such as diabetes) or overproduction (such as hyperthyroidism) of a hormone or other endogenous biologically active substance. Use of the transposon-based vectors of the present invention to integrate nucleotide sequences encoding RNA molecules, such as anti-sense RNA or short interfering RNA, is also contemplated as a part of the present 30 invention.

Additionally, the transposon-based vectors of the present invention may be used to provide cells or tissues with "beacons", such as receptor molecules, for binding of therapeutic agents in order to provide tissue and cell specificity for the

therapeutic agents. Several promoters and exogenous genes can be combined in one vector to produce progressive, controlled treatments from a single vector delivery.

The following examples will serve to further illustrate the present invention without, at the same time, however, constituting any limitation thereof. On the contrary, it is to be clearly understood that resort may be had to various embodiments, modifications and equivalents thereof which, after reading the description herein, may suggest themselves to those skilled in the art without departing from the spirit of the invention.

10

EXAMPLE 1

Preparation of Transposon-Based Vector pTnMod

A vector was designed for inserting a desired coding sequence into the genome of eukaryotic cells, given below as SEQ ID NO:1. The vector of SEQ ID NO:1, termed pTnMod, was constructed and its sequence verified.

15

This vector employed a cytomegalovirus (CMV) promoter. A modified Kozak sequence (ACCATG) (SEQ ID NO:13) was added to the promoter. The nucleotide in the wobble position in nucleotide triplet codons encoding the first 10 amino acids of transposase was changed to an adenine (A) or thymine (T), which did not alter the amino acid encoded by this codon. Two stop codons were added and a synthetic polyA was used to provide a strong termination sequence. This vector uses a promoter designed to be active soon after entering the cell (without any induction) to increase the likelihood of stable integration. The additional stop codons and synthetic polyA insures proper termination without read through to potential genes downstream.

25

The first step in constructing this vector was to modify the transposase to have the desired changes. Modifications to the transposase were accomplished with the primers High Efficiency forward primer (Hef) Altered transposase (ATS)-Hef 5' ATCTCGAGACCATGTGTGAACCTTGATATTACATGATTCTCTTTACC 3' (SEQ ID NO:10) and Altered transposase- High efficiency reverse primer (Her) 5' 30 GATTGATCATTATCATAATTCCCCAAAGCGTAACC 3' (SEQ ID NO:11, a reverse complement primer). In the 5' forward primer ATS-Hef, the sequence CTCGAG (SEQ ID NO:12) is the recognition site for the restriction enzyme Xho I, which permits directional cloning of the amplified gene. The sequence ACCATG

(SEQ ID NO:13) contains the Kozak sequence and start codon for the transposase and the underlined bases represent changes in the wobble position to an A or T of codons for the first 10 amino acids (without changing the amino acid coded by the codon). Primer ATS-Her (SEQ ID NO:11) contains an additional stop codon TAA in addition

5 to native stop codon TGA and adds a Bcl I restriction site, TGATCA (SEQ ID NO:14), to allow directional cloning. These primers were used in a PCR reaction with pTnLac (p defines plasmid, tn defines transposon, and lac defines the beta fragment of the lactose gene, which contains a multiple cloning site) as the template for the transposase and a FailSafe™ PCR System (which includes enzyme, buffers,

10 dNTP's, MgCl₂ and PCR Enhancer; Epicentre Technologies, Madison, WI). Amplified PCR product was electrophoresed on a 1% agarose gel, stained with ethidium bromide, and visualized on an ultraviolet transilluminator. A band corresponding to the expected size was excised from the gel and purified from the agarose using a Zymo Clean Gel Recovery Kit (Zymo Research, Orange, CA).

15 Purified DNA was digested with restriction enzymes Xho I (5') and Bcl I (3') (New England Biolabs, Beverly, MA) according to the manufacturer's protocol. Digested DNA was purified from restriction enzymes using a Zymo DNA Clean and Concentrator kit (Zymo Research).

Plasmid gWhiz (Gene Therapy Systems, San Diego, CA) was digested with

20 restriction enzymes Sal I and BamH I (New England Biolabs), which are compatible with Xho I and Bcl I, but destroy the restriction sites. Digested gWhiz was separated on an agarose gel, the desired band excised and purified as described above. Cutting the vector in this manner facilitated directional cloning of the modified transposase (mATS) between the CMV promoter and synthetic polyA.

25 To insert the mATS between the CMV promoter and synthetic polyA in gWhiz, a Stratagene T4 Ligase Kit (Stratagene, Inc. La Jolla, CA) was used and the ligation set up according to the manufacturer's protocol. Ligated product was transformed into *E. coli* Top10 competent cells (Invitrogen Life Technologies, Carlsbad, CA) using chemical transformation according to Invitrogen's protocol.

30 Transformed bacteria were incubated in 1 ml of SOC (GIBCO BRL, CAT# 15544-042) medium for 1 hour at 37° C before being spread to LB (Luria-Bertani media (broth or agar)) plates supplemented with 100 µg/ml ampicillin (LB/amp plates). These plates were incubated overnight at 37° C and resulting colonies picked to

LB/amp broth for overnight growth at 37° C. Plasmid DNA was isolated using a modified alkaline lysis protocol (Sambrook et al., 1989), electrophoresed on a 1% agarose gel, and visualized on a U.V. transilluminator after ethidium bromide staining. Colonies producing a plasmid of the expected size (approximately 6.4 kbp) 5 were cultured in at least 250 ml of LB/amp broth and plasmid DNA harvested using a Qiagen Maxi-Prep Kit (column purification) according to the manufacturer's protocol (Qiagen, Inc., Chatsworth, CA). Column purified DNA was used as template for sequencing to verify the changes made in the transposase were the desired changes and no further changes or mutations occurred due to PCR amplification. For 10 sequencing, Perkin-Elmer's Big Dye Sequencing Kit was used. All samples were sent to the Gene Probes and Expression Laboratory (LSU School of Veterinary Medicine) for sequencing on a Perkin-Elmer Model 377 Automated Sequencer.

Once a clone was identified that contained the desired mATS in the correct orientation, primers CMVf-NgoM IV (5' TTGCCGGCCATCAGATTGGCTAT (SEQ 15 ID NO:15); underlined bases denote NgoM IV recognition site) and Syn-polyA-BstE II (5' AGAGGTCACCGGGTCAATTCTTCAGCACCTGGTA (SEQ ID NO:16); underlined bases denote BstE II recognition site) were used to PCR amplify the entire CMV promoter, mATS, and synthetic polyA for cloning upstream of the transposon in pTnLac. The PCR was conducted with FailSafe™ as described above, purified 20 using the Zymo Clean and Concentrator kit, the ends digested with NgoM IV and BstE II (New England Biolabs), purified with the Zymo kit again and cloned upstream of the transposon in pTnLac as described below.

Plasmid pTnLac was digested with NgoM IV and BstE II to remove the ptac 25 promoter and transposase and the fragments separated on an agarose gel. The band corresponding to the vector and transposon was excised, purified from the agarose, and dephosphorylated with calf intestinal alkaline phosphatase (New England Biolabs) to prevent self-annealing. The enzyme was removed from the vector using a Zymo DNA Clean and Concentrator-5. The purified vector and CMVp/mATS/polyA 30 were ligated together using a Stratagene T4 Ligase Kit and transformed into *E. coli* as described above.

Colonies resulting from this transformation were screened (mini-preps) as described above and clones that were the correct size were verified by DNA sequence

analysis as described above. The vector was given the name pTnMod (SEQ ID NO:1) and includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to base pairs 1-130 of pBluescriptII sk(-).

5 Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 133 -1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems), corresponding to bp 229-1873 of pGWiz. The CMV promoter was modified by the addition of an ACC sequence upstream of ATG.

10 Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons for stability of the transposase mRNA and for the expression of protein. More specifically, in each of the 15 codons for the first ten amino acids of the transposase, G or C was changed to A or T when such a substitution would not alter the amino acid that was encoded.

Base pairs 2988-2993 are two engineered stop codons.

Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

20 Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of 10 pGWiz.

Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from pNK2859.

25 Base pairs 3762 - 3831 are the 70 bp of the left insertion sequence recognized by the transposon Tn10.

Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

30 Base pairs 3838 - 4527 are the multiple cloning site from pBluescriptII sk(20), corresponding to bp 924-235 of pBluescriptII sk(-). This multiple cloning site may be used to insert any coding sequence of interest into the vector.

Base pairs 4528-4532 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4533 - 4602 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 4603 - 4644 are non-coding λ DNA that is residual from pNK2859.

Base pairs 4645 - 5488 are non-coding DNA that is residual from pNK2859.

5 Base pairs 5489 - 7689 are from the pBluescriptII SK(-) base vector (Stratagene, Inc.), corresponding to bp 761-2961 of pBluescriptII SK(-).

Completing pTnMod is a pBlueScript backbone that contains a colE I origin of replication and an antibiotic resistance marker (ampicillin).

It should be noted that all non-coding DNA sequences described above can be 10 replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

15 All plasmid DNA was isolated by standard procedures. Briefly, *Escherichia coli* containing the plasmid was grown in 500 mL aliquots of LB broth (supplemented with an appropriate antibiotic) at 37°C overnight with shaking. Plasmid DNA was recovered from the bacteria using a Qiagen Maxi-Prep kit (Qiagen, Inc., Chatsworth, CA) according to the manufacturer's protocol. Plasmid DNA was resuspended in 500 μL of PCR-grade water and stored at -20°C until used.

EXAMPLE 2

20 *Preparation of Transposon-Based Vector pTnMod (CMV/Red)*

A vector was designed for inserting a reporter gene (DsRed) under the control of the CMV promoter into the genome of vertebrate cells given below as SEQ ID NO:2. The reporter gene chosen was the DsRed gene, driven by the immediate early cytomegalovirus promoter, to produce a plasmid called pTnCMV/DsRed. The DsRed 25 gene product is a red fluorescent protein from an IndoPacific sea anemone, Discosoma sp., which fluoresces bright red at 558 nm. It is to be understood that the reporter gene, i.e., the DsRed gene, is only one embodiment of the present invention and that any gene of interest may be inserted into the plasmid in place of the DsRed reporter gene in any Experiment described herein.

30 The vector of SEQ ID NO:2, named pTnMod (CMV/Red), was constructed, and its sequence verified by re-sequencing. SEQ ID NO:2, pTnMod (CMV/Red), includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to bp 1-130 of pBluescriptII sk(-).

Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

5 Base pairs 133 -1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems, corresponding to bp 229-1873 of pGWiz.

Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

10 Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons as discussed above.

Base pairs 2988-2993 are two engineered stop codons.

Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

15 Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of pGWiz.

Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from pNK2859.

20 Base pairs 3762 - 3831 are the 70 bp of the left insertion sequence recognized by the transposon Tn10.

Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 3838 - 4044 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 924-718 of pBluescriptII sk(-).

25 Base pairs 4045-4048 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4049-5693 are the CMV promoter/enhancer, taken from vector pGWiz (Gene Therapy Systems), corresponding to bp 229-1873 of pGWiz.

30 Base pairs 5694-5701 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 5702 - 6617 are the DsRed reporter coding sequence, including polyA sequence, from pDsRed1.1 (Clontech), corresponding to bp 77 - 992 of pDsRed1.1.

Base pairs 6618 - 7101 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 718-235 of pBluescriptII sk(-).

Base pairs 7102-7106 are a residue from ligation of restriction enzyme sites used in constructing the vector.

5 Base pairs 7107 - 7176 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 7177 - 7218 are non-coding λ DNA that is residual from pNK2859.

Base pairs 7219 - 8062 are non-coding DNA that is residual from pNK2859.

10 Base pairs 8063 - 10263 are from the pBluescriptII sk(-) base vector (Stratagene, Inc.), corresponding to bp 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s).

EXAMPLE 3

15 *Preparation of Transposon-Based Vector pTnMod (Oval/Red) - Chicken*

A vector was designed for inserting a reporter gene (DsRed) under the control of the ovalbumin promoter, and including the ovalbumin signal sequence, into the genome of a bird. One version of this vector is given below as SEQ ID NO:3. The vector of SEQ ID NO:3, named pTnMod (Oval/Red) - Chicken, includes chicken 20 ovalbumin promoter and signal sequences.

SEQ ID NO:3, pTnMod (Oval/Red) - Chicken, includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to bp 1-130 of pBluescriptII sk(-).

25 Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 133 -1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems, corresponding to bp 229-1873 of pGWiz.

20 Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons as discussed above.

Base pairs 2988-2993 are two engineered stop codons.

Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of pGWiz.

5 Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from 10 pNK2859.

10 Base pairs 3762 - 3831 are the 70 bp of the left insertion sequence recognized by the transposon Tn10.

Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 3838 - 4044 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 924-718 of pBluescriptII sk(-).

15 Base pairs 4045-4049 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4050 - 4951 contain upstream elements of the (including SDRE, steroid-dependent response element). See GenBank accession number J00895 M24999, bp 431-1332. Base pairs 4952-4959 are a residue from ligation of 20 restriction enzyme sites used in constructing the vector.

Base pairs 4960 - 5112 are the chicken ovalbumin signal sequence (GenBank accession number J00895 M24999, bp 2996-3148).

Base pairs 5113-5118 are a residue from ligation of restriction enzyme sites used in constructing the vector.

25 Base pairs 5119 - 6011 are the DsRed reporter coding sequence, including polyA sequence, from pDsRed1.1 (Clontech), corresponding to bp 100 - 992 of pDsRed1.1.

Base pairs 6012-6017 are a residue from ligation of restriction enzyme sites used in constructing the vector.

30 Base pairs 6018 - 6056 are part of the multiple cloning site of the ZeroBlunt Topo cloning vector (Invitrogen), corresponding to bp 337-377 of ZeroBlunt.

Base pairs 6057-6062 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 6063 - 6495 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 667-235 of pBluescriptII sk(-).

Base pairs 6496-6500 are a residue from ligation of restriction enzyme sites used in constructing the vector.

5 Base pairs 6501 - 6570 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 6571 - 6612 are non-coding λ DNA that is residual from pNK2859.

Base pairs 6613 - 7477 are non-coding DNA that is residual from pNK2859.

10 Base pairs 7478 - 9678 are from the pBluescriptII sk(-) base vector (Stratagene, Inc.), corresponding to bp 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s).

EXAMPLE 4

15 *Preparation of Transposon-Based Vector pTnMod(Oval/Red) - Quail*

A vector was designed for inserting a reporter gene (DsRed) under the control of the ovalbumin promoter, and including the ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:4. The vector of SEQ ID NO:4, named pTnMod (Oval/Red) - Quail, has been constructed, and selected portions of the 20 sequence have been verified by re-sequencing.

SEQ ID NO:4, pTnMod (Oval/Red) - Quail, includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to bp 1-130 of pBluescriptII sk(-).

25 Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 133 - 1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems), corresponding to bp 229-1873 of pGWiz.

30 Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons as discussed above.

Base pairs 2988-2993 are two engineered stop codons. Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of pGWiz.

5 Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from pNK2859.

Base pairs 3762 - 3831 are the 70 base pairs of the left insertion sequence recognized by the transposon Tn10.

10 Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 3838 - 4044 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 924-718 of pBluescriptII sk(-).

15 Base pairs 4045-4049 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4050 - 4934 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332). Some deletions 20 were noted in the quail sequence, as compared to the chicken sequence.

Base pairs 4935-4942 are a residue from ligation of restriction enzyme sites used in constructing the vector.

25 Base pairs 4943 - 5092 are the Japanese quail ovalbumin signal sequence. The quail signal sequence was isolated by its high degree of homology to the chicken signal sequence (GenBank accession number J00895 M24999, base pairs 2996-3148). Some deletions were noted in the quail sequence, as compared to the chicken sequence.

Base pairs 5093-5098 are a residue from ligation of restriction enzyme sites used in constructing the vector.

30 Base pairs 5099 - 5991 are the DsRed reporter coding sequence, including polyA sequence, from pDsRed1.1 (Clontech), corresponding to bp 100 - 992 of pDsRed 1.1.

Base pairs 5992-5997 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 5998 - 6036 are part of the multiple cloning site of the ZeroBlunt Topo cloning vector (Invitrogen), corresponding to base pairs 337-377 of ZeroBlunt.

5 Base pairs 6037-6042 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 6043 - 6475 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 667-235 of pBluescriptII sk(-).

10 Base pairs 6476-6480 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 6481 - 6550 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 6551 - 6592 are non-coding λ DNA that is residual from pNK2859.

Base pairs 6593 - 7457 are non-coding DNA that is residual from pNK2859.

15 Base pairs 7458 - 9658 are from the pBluescriptII sk(-) base vector (Stratagene, Inc.), corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s).

20

EXAMPLE 5

Transfection of Stage X Japanese Quail Eggs with pTnMod(Oval/Red) – Quail via embryo injection

Transgenic Japanese quail were produced by transfecting Stage X embryos and the heritability of the transgene delivered by embryo transfection was established.

25 More specifically, fertile eggs were collected in the morning and placed at 15° C until enough were collected for injection, but were held no longer than 7 days. Stage X embryos (eggs) were assigned to one of two treatment groups. Before treatment, each egg was incubated on its side at room temperature for about 2 hours to allow the embryo to move to "top dead center" (TDC). Each egg was transfected by drilling a 1
30 mm hole (directly above the embryo) through the shell without penetrating the underlying shell membrane. A 0.5 ml syringe fitted with a 28 gauge needle was used to deliver DNA complexed to a transfecting reagent, i.e. SUPERFECT®, in a 50 µl volume. An adhesive disc was used to seal the hole and provide a label for treatment identification. After all eggs were transfected, they were set in an incubator with the
35 adhesive disc pointing upward for hatching.

Each bird that hatched was bled at one week of age, DNA was extracted from blood cells, and PCR was conducted using 28s primers as a positive control and primers specific to DsRed. Any bird that was negative was terminated, while positive birds were monitored to determine maintenance of the transgene. Birds consistently positive were maintained until sexual maturity and bred. Positive male and female birds were mated. The eggs of mated hens were hatched and the resulting chicks, the G1 generation, were evaluated to determine if they were transgenic. All G1s resulting from this mating were bled and PCR conducted as described above.

5 Egg injection: Two treatment groups and one control group were used for this experiment. Vector pTnMod (Oval/Red) in supercoiled form (Treatment 1) and in linear form (Treatment 2) were used to transfect 15 eggs per treatment. To obtain linear DNA for this experiment, pTnMod (Oval/Red) was digested with NgoM IV, column purified, and resuspended in TE buffer.

10 15 Each egg was injected with 0.75 µg of DNA complexed with SUPERFECT® in a 1:3 ratio in a total injection volume of 50 µl Hank's Balanced Salt solution (HBSS) was used to bring the volume to 50 µl. The DNA Superfect mixture must be allowed to incubate (for complex formation) at room temperature for 10 minutes prior to injection and must be used within 40 minutes post initial mixing. Eggs were incubated as described above after injection.

20 25 Results: In the supercoiled injection group, 2 females and 1 male were identified as PCR positive using primers specific to the DsRed coding sequence. These birds were mated as described above. Blood was taken from the G1 chicks and PCR was conducted. The results showed that the transgene was incorporated into the gametes of these birds. The G1 chicks from these birds were examined on a weekly basis until it was verified that the gene was not present or enough transgenic G1s were obtained to initiate a breeding flock of fully transgenic birds. Eggs from these G1 chicks expressed DsRed protein in the albumin portion of their eggs.

EXAMPLE 6

30 *Intratesticular Injection of Chickens with pTnMod(CMV/Red) (SEQ ID NO:2)*

Immature birds of different ages (4, 6, 8, 10, 12, and 14 weeks) were placed under anesthesia and injected in the testes with the construct pTnMod(CMV/Red). A saline solution containing 1-5 µg of purified DNA vector, mixed with SUPERFECT® transfecting reagent (Qiagen, Valencia, CA) in a 1:6 (wt:vol) ratio. The volume of 35 saline was adjusted so that the total volume injected into each testis was 150-200 µl, depending on the age and size of the bird. For the 4- and 6-week-old chickens, 1 µg DNA in 150 µl was injected in each testis, divided into three doses of 50 µl each. For

the older birds, 200 μ l total volume was injected, containing either 3 μ g DNA (for 8-week-old birds) or 5 μ g DNA (for older birds) per testis. First, one testis was surgically exposed prior to injection. After injection, the incision was sutured, and the sequence was repeated for the alternate testis.

5 From six to nine months post-surgery, weekly sperm samples were taken from each injected bird, as well as from control birds. Each sperm sample was evaluated for uptake and expression of the injected gene. Samples were evaluated by PCR on whole sperm, within one week after collection.

10 Approximately 100 male white leghorn chickens, in groups of 5-26, at ages 4, 6, 8, 10, 12, and 14 weeks, were used as this is the age range in which it is expected that the testes are likely to be most "receptive." In this age range, the blood/testis barrier has not yet formed, and there is a relatively high number of spermatogonia relative to the numbers of other cell types, e.g., spermatids, etc. See J. Kumaran et al., 1949. Poultry Sci., vol. 29, pp. 511-520. See also E. Oakberg, 1956. Am. J. Anatomy, 15 vol. 99, pp. 507-515; and P. Kluin et al., 1984. Anat. EmbryoL, vol. 169, pp. 73-78.

15 The experimental and control males were obtained from commercial sources at one day of age, and maintained in brooders until used. The male birds were housed in temperature-controlled spaces in individual standard caging as they approached maturity. They were given water and standard commercial feed ad lib. They were 20 kept initially in a 23:1 hour light/dark cycle, stepped down at approximately weekly intervals to a 15:8 hour light/dark cycle, as this regimen has been reported to optimize sexual maturity and fertility.

Surgical and DNA Injection Procedures

25 At the appropriate ages, groups of individual males were starved overnight and then subjected to transgene delivery by direct intratesticular injection of DNA by experienced animal surgeons. Each male was anesthetized with isoflurane via a simplified gas machine.

30 Various devices and anesthesia machines have previously been described for administering isoflurane (and other gaseous anesthetics) to birds. See Alsage et al., Poultry Sci., 50:1876-1878 (1971); Greenlees et al., Am. J. Vet. Res., vol. 51, pp. 757-758 (1990). However, these prior techniques are somewhat cumbersome and complex to implement. A novel and much simpler system to administer isoflurane (or other gaseous) anesthesia was developed due to the deficiencies in the prior art, a

system that we found worked well on all ages of chicks. A standard nose cone was placed over the chick's head, similar to the system that has been used for decades to administer ether to mice. A plastic tube approximately 3.5 cm in diameter and 12 cm long was filled with cotton, into which was poured approximately 2 mL isoflurane 5 (Abbott Laboratories, Chicago). The chick's head was placed partially into the cylinder, and was held in place there intermittently throughout the surgery as required to maintain the proper plane of anesthesia, without overdosing.

Each anesthetized bird was positioned on its side on an animal board with cords tractioning the wings and feet to allow access to the testes area. The area was 10 swabbed with 0.5% chlorhexidine, and a 2 cm dorsolateral incision was made in the skin over the testis (similar to the procedure commonly used for caponization). A small-animal retractor was used to spread the last two ribs, exposing the testis. The DNA solution was then mixed with SUPERFECT® (Qiagen) according to the manufacturer's protocol, approximately a 1:6 wt/vol ratio, to a final concentration of 15 0.01 - 0.05 µg/µl. This resulted in 1 - 5 µg total DNA (in a 150-200 µl volume) being injected into each testis, spread over three injection sites: one at each end of the testis, and one in the middle.

The injection device was a standard 25 gauge, 1/2 inch (1.27 cm) hypodermic needle, attached to a 50, 100, or 200 µl syringe. Approximately 5 mm of the needle 20 tip was bent at a 90 degree angle, to facilitate insertion into the testes. Approximately 50 - 70 µl of the DNA-SUPERFECT® solution was injected into each of three sites per testis. The multiple injections were calculated to suffuse the DNA throughout the whole testis, the idea being to promote contact between DNA and spermatogonia as much as feasible. We estimated that our procedure resulted in the injection of about 25 100,000 DNA molecules per spermatogonium. The construct used in these tests was a highly potent constitutive modified CMV promoter, operatively linked to the dsRed gene as shown in SEQ ID NO:2.

Following injection, the incision was closed in two layers with 4-0 absorbable suture, and then the contralateral testis was similarly exposed and injected. Following 30 surgery, each bird was returned to its cage to recover. One hundred thirteen males were ultimately used in the experimental regimen to increase the overall likelihood of success, along with 4 control birds (16 weeks 20 old) subjected to sham surgery (with injections containing only the transfection reagent.

Evaluation of Birds

Thus, a total of 113 white leghorn chickens were injected with the DNA vector in groups of 5-26 at varying ages. Fourteen birds were transformed at 4 weeks, 23 birds at 6 weeks; 26 birds at 8 weeks; 23 birds at 10 weeks; 5 birds at 12 weeks; and 5 22 birds at 14 weeks. Sixteen birds died before they could be sampled, so to date, 97 roosters have been sampled, plus the four controls. Birds were evaluated at 18-24 weeks of age for (a) potential transformation in the sperm, and (b) successful testis transfection. Sperm samples were obtained from each rooster by manual manipulation using standard techniques. The sperm were washed, and their DNA was 10 extracted following the techniques of G. Mann et al., 1993. J. Reprod. Fert., 99:505-12. The samples were then frozen until analyzed. Evaluation was conducted by PCR analysis to detect DNA integration into the sperm, or into any of the testicular cells. Additionally, selected testes were harvested at the end of the sperm sampling period.

Of 97 birds tested, at least 22 showed probable positive results. Positive 15 results were observed at all transformation ages, except for 4 weeks, which was not tested. At least two birds were confirmed positive by PCR of sperm, conducted four months after the initial injection. These results were transient in many cases, however since it was believed that the DsRed gene product used in these initial proof of concept experiments was toxic. Nevertheless, the positive PCR results presumptively 20 demonstrated that the transgene was incorporated into spermatogonia (before puberty), and that it was carried in transgenic sperm. Such sperm could then transmit the gene to subsequent generations, resulting in the production of true, germ-line transgenic "founder" birds.

To further confirm that the DNA had been incorporated into the sperm, and 25 that contaminating vector was not being detected from other sources, it was confirmed through PCR on sperm of experimental birds, and on positive and negative controls that the sperm of the experimental birds lacked DNA encoding the transposase. The design of the preferred transposon-based vector is such that the sequence encoding the transposase is contained in the vector, but is not incorporated into the transformed 30 chromosome. Thus, presence of the exogenous coding sequence, coupled with absence of the transposase gene, is strong evidence for incorporation of the exogenous coding sequence, or transgene.

These results demonstrated proof of concept, as positive PCR results were obtained from the sperm of treated birds. Interpretation of these preliminary results was made more difficult by the fact that the modified CMV promoter used in the experiment was probably too "hot." As the DsRed product is not secreted from the 5 cells, the product built up intracellularly to levels that were toxic, frequently killing the cells. Even this result, of course, means that the transformation was successful. The transgene could not have killed the cells otherwise.

In order to resolve to the problem with toxicity of the DsRed gene product, experiments were conducted using a different reporter gene operably linked to the 10 ovalbumin promoter, so that the transgene was expressed in the egg white. These experiments are provided in Examples 12-15 below.

EXAMPLE 7

Transfection of Male White Leghorn Chickens Using the Vector pTnMod(Oval/Red) – 15 Quail (SEQ ID NO:4) via Testicular Injections

In further experiments conducted on leghorn chickens, it was demonstrated that chickens injected intratesticularly at 8, 10, 12, or 14 weeks of age, had, on average, approximately 40% positive sperm between 6 and 8 months after injection. In other experiments, successful transfection was achieved with chickens injected at 20 13 weeks of age.

Forty-nine white leghorn roosters approximately 8, 10, 12, or 14 weeks of age were obtained and housed. Birds were identified, wing banded, and assigned to a treatment group. If appropriate (based on testes size and vascularization), one testis was caponized and the entire DNA injection volume was delivered to the remaining 25 testis. Thirty-two males received DNA injections of 5 μ g DNA/testis at a 1:3 ratio of DNA to SUPERFECT®. The remaining birds were used as controls. After injection, all birds were mated with at least 5 females and observed until sexual maturity and egg-laying began. All eggs collected prior to peak egg production (approximately 24 weeks of age for the hens) were incubated and candled to determine embryo presence. 30 Any embryos identified were incubated to hatch to extract DNA, PCR was conducted, and transgene presence was determined.

Roosters positive for the pTnMod(Oval/Red) – Quail construct were kept to produce F1 offspring (eggs collected at peak production). Offspring from this hatch

were bled, DNA extracted from the blood, and PCR conducted using primers specific for the DsRed gene. It was determined that 77% of the offspring were transgenic.

EXAMPLE 8

5 *Transfection of Mature Male Japanese Quail using the vector pTnMod(Oval/Red) – Quail (SEQ ID NO:4) via Testicular Injections*

Twelve sexually mature males (at approximately 13 weeks of age) underwent surgery for testicular injection as described above for chickens. At 21-28 days of age, the birds were identified, leg banded, debeaked, and separated based on sex.

10 Injections comprised 5 µg/testes of the vector in concentrations 1:3 or 1:10 for SUPERFECT® or a 1:1 ratio with Mirrus. The study consisted of 3 treatment groups with 5 males in the 1:3 DNA:SUPERFECT® group, 3 males in the 1:10 DNA:SUPERFECT® group, and 4 males in the 1:1 Mirrus group. All surgeries were conducted in one day.

15 Any unincorporated DNA was allowed to clear from the testes by holding the birds for 19 days before mating with females. At 15 weeks of age, 2 age-matched females were housed with each treated male. The presence of the transfected DNA was determined in the fertilized eggs during the second week of egg lay. The subsequent eggs collected from parents producing positively identified transgenic 20 eggs were collected and stored until taken to hatch.

PCR performed on the sperm of quail injected at three months of age indicated successful incorporation of the DsRed transgene into the quail sperm.

EXAMPLE 9

25 *Transfection of Immature Male Japanese Quail using the vector pTnMod(Oval/Red) – Quail (SEQ ID NO:4) via Testicular Injections*

Approximately 450 quail eggs were set and hatched. At 21-28 days of age, the birds were identified, wingbanded, debeaked, and separated based on sex. At 4 weeks of age, 65 male birds underwent surgery and testicular injections as described above.

30 Injections comprised a control and 2 µg/testes of the vector in varying concentrations (0, 1/3, 1/5, and 1/10) of three different transfection reagents: 1) SUPERFECT®, 2) Mirus/Panvera and 3) Doper. The study comprised 13 treatment groups with 5 males per group. One transfection reagent was administered per day.

At 7 weeks of age, 2 age-matched females were housed with each treated male. The presence of the transfected DNA was determined in the fertilized eggs during the second week of egg lay. The subsequent eggs collected from parents producing positively identified transgenic eggs were collected and stored until taken 5 to hatch. PCR performed on the sperm of quail injected at four and five weeks of age indicated successful incorporation of the DsRed transgene into the quail sperm.

EXAMPLE 10

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/p146/PA) –

10 Chicken

A vector is designed for inserting a p146 gene under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:29.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) 15 (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 – 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 – 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

20 Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector pNK2859.

25 Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

30 Base pairs 4050 - 4951 are a chicken ovalbumin promoter (including SDRE) that corresponds to base pairs 431-1332 of the chicken ovalbumin promoter in GenBank Accession Number J00895 M24999.

Base pairs 4958 - 6115 are a chicken ovalbumin signal sequence and Ovalbumin gene that correspond to base pairs 66-1223 of GenBank Accession Number V00383.1 (The STOP codon being omitted).

5 Base pairs 6122 - 6271 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

Base pairs 6272 - 6316 are a p146 sequence (synthetic) with 2 added stop codons.

Base pairs 6324 - 6676 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

10 Base pairs 6682 - 7114 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

Base pairs 7120- 7189 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7190 - 7231 are λ DNA that is residual from pNK2859.

15 Base pairs 7232 - 8096 are non coding DNA that is residual from pNK2859.

Base pairs 8097 - 10297 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

20 It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

EXAMPLE 11

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/p146/PA) - Quail

25 A vector is designed for inserting a p146 gene under the control of a quail ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:30.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

30 Base pairs 133 - 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 - 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922-2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector pNK2859.

5 Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

10 Base pairs 4050 - 4938 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332).

15 Bp 4945 - 6092 are a quail ovalbumin signal sequence and ovalbumin gene that corresponds to base pairs 54 – 1201 of GenBank accession number X53964.1. (The STOP codon being omitted).

Base pairs 6097 - 6246 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

20 Base pairs 6247 – 6291 are a p146 sequence (synthetic) with 2 added stop codons.

Base pairs 6299 – 6651 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6657 - 7089 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

25 Base pairs 7095- 7164 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7165 - 7206 are λ DNA that is residual from pNK2859.

Base pairs 7207 – 8071 are non coding DNA that is residual from pNK2859.

30 Base pairs 8072 - 10272 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

5

EXAMPLE 12

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/ProIns/PA) – Chicken

A vector is designed for inserting a proinsulin gene under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal 10 sequence, into the genome of a bird given below as SEQ ID NO:31.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 – 1777 are a CMV promoter/enhancer taken from vector 15 pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 – 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy 20 Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector pNK2859.

Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence 25 (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

Base pairs 4050 - 4951 are a chicken ovalbumin promoter (including SDRE) that corresponds to base pairs 431-1332 of the chicken ovalbumin promoter in 30 GenBank Accession Number J00895 M24999.

Base pairs 4958 - 6115 are a chicken ovalbumin signal sequence and ovalbumin gene that correspond to base pairs 66-1223 of GenBank Accession Number V00383.1. (The STOP codon being omitted).

Base pairs 6122 - 6271 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

Base pairs 6272 – 6531 are a proinsulin gene.

Base pairs 6539 – 6891 are a synthetic polyadenylation sequence from pGWiz 5 (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6897 - 7329 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

Base pairs 7335- 7404 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

10 Base pairs 7405 - 7446 are λ DNA that is residual from pNK2859.

Base pairs 7447 – 8311 are non coding DNA that is residual from pNK2859.

Base pairs 8312 - 10512 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

15 It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

EXAMPLE 13

20 *Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/ProIns/PA) – Quail*

A vector is designed for inserting a proinsulin gene under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:32.

25

Base pairs 1 -130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 – 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

30

Base pairs 1780 – 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector pNK2859.

5 Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

10 Base pairs 4050 - 4938 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332). Some deletions were noted in the quail sequence, as compared to the chicken sequence.

15 Base pairs 4945 - 6092 are a quail ovalbumin signal sequence and ovalbumin gene that corresponds to base pairs 54 – 1201 of GenBank accession number X53964.1. (The STOP codon being omitted).

Base pairs 6093 - 6246 are a TAG sequence containing a gp41 hairpin loop from HIV I an enterokinase cleavage site and a spacer (synthetic).

20 Base pairs 6247 – 6507 are a proinsulin gene.

Base pairs 6514 – 6866 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6867 - 7303 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

25 Base pairs 7304- 7379 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7380 - 7421 are λ DNA that is residual from pNK2859.

Base pairs 7422 – 8286 are non coding DNA that is residual from pNK2859.

Base pairs 8287 - 10487 are pBlueScript sk(-) base vector (Stratagene, Inc.)

30 corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

5

EXAMPLE 14

Transfection of Immature Leghorn Roosters using a Transposon-based Vector containing a Proinsulin Gene via Testicular Injections

Vectors containing the elements Oval promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:31) and CMV promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:42) were each injected into the testes of 11 week old white leghorn roosters. These birds were held under normal conditions until sexual maturity was reached.

At the time of sexual maturity, each bird was handled and manipulated to obtain sperm. Sperm samples were collected in Hank's Buffered Salt Solution (HBSS) and stored at either -20° C or 4° C until needed. DNA was extracted from sperm using a MoBio Ultra Clean DNA Bloodspin Kit (MoBio laboratories, Solana Beach CA). Fifty microliters of sperm was used in the DNA extraction protocol and the purified genomic DNA eluted in 100 µl of water. In each PCR reaction, approximately 0.5 – 0.75 µg of genomic DNA was used with primers anchored in the entag-1 (5') and the synthetic polyA-2 (3'), which amplify a 685 bp fragment. Five of nine birds gave positive reactions for the presence of the appropriate vector construct. These birds were then mated with normal females.

Birds that did not yield positive results with PCR on the sperm were sacrificed, their testes removed, and DNA extracted using an approximately 25 mg piece of tissue in a Qiagen DNEasy Tissue Kit; purified DNA was eluted in 200 µl water and PCR conducted as described above. Two of these birds gave a very strong, positive PCR reaction.

EXAMPLE 15

30 *Transfection of Japanese Quail using a Transposon-based Vector containing a Proinsulin Gene via Oviduct Injections*

Two experiments were conducted in Japanese quail using transposon-based vectors containing either Oval promoter/Oval gene/GP41 Enterokinase

TAG/Proinsulin/Poly A (SEQ ID NO:31) or CMV promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:42).

In the first experiment, the Oval promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A containing construct was injected into the oviduct of sexually mature quail; three hens received 5 µg at a 1:3 Superfect ratio and three received 10 µg at a 1:3 Superfect ratio. As of the writing of the present application, at least one bird that received 10 µg of DNA was producing human proinsulin in egg white (other birds remain to be tested). This experiment indicates that 1) the DNA has been stable for at least 3 months; 2) protein levels are comparable to those observed with a constitutive promoter such as the CMV promoter; and 3) sexually mature birds can be injected and results obtained without the need for cell culture.

In the second experiment, the transposon-based vector containing CMV promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A was injected into the oviduct of sexually immature Japanese quail. A total of 9 birds were injected. Of the 8 survivors, 3 produced human proinsulin in the white of their eggs for over 6 weeks. An ELISA assay described in detail below was developed to detect GP41 in the fusion peptide (Oval gene/GP41 Enterokinase TAG/Proinsulin) since the GP41 peptide sequence is unique and not found as part of normal egg white protein. In all ELISA assays, the same birds produced positive results and all controls worked as expected.

ELISA Procedure: Individual egg white samples were diluted in sodium carbonate buffer, pH 9.6, and added to individual wells of 96 well microtiter ELISA plates at a total volume of 0.1 ml. These plates were then allowed to coat overnight at 4°C. Prior to ELISA development, the plates were allowed warm to room temperature. Upon decanting the coating solutions and blotting away any excess, non-specific binding of antibodies was blocked by adding a solution of phosphate buffered saline (PBS), 1% (w/v) BSA, and 0.05% (v/v) Tween 20 and allowing it to incubate with shaking for a minimum of 45 minutes. This blocking solution was subsequently decanted and replaced with a solution of the primary antibody (Goat Anti-GP41 TAG) diluted in fresh PBS/BSA/Tween 20. After a two hour period of incubation with the primary antibody, each plate was washed with a solution of PBS and 0.05% Tween 20 in an automated plate washer to remove unbound antibody. Next, the secondary antibody, Rabbit anti-Goat Alkaline Phosphatase-conjugated, was diluted in PBS/BSA/Tween 20 and allowed to incubate 1 hour. The plates were then

subjected to a second wash with PBS/Tween 20. Antigen was detected using a solution of *p*-Nitrophenyl Phosphate in Diethanolamine Substrate Buffer for Alkaline Phosphatase and measuring the absorbance at 30 minutes and 1 hour.

5

EXAMPLE 16

Optimization of Intra-oviduct and Intra-ovarian Arterial Injections

Overall transfection rates of oviduct cells in a flock of chicken or quail hens are enhanced by synchronizing the development of the oviduct and ovary within the flock. When the development of the oviducts and ovaries are uniform across a group

10 of hens and when the stage of oviduct and ovarian development can be determined or predicted, timing of injections is optimized to transfect the greatest number of cells. Accordingly, oviduct development is synchronized as described below to ensure that a large and uniform proportion of oviduct secretory cells are transfected with the gene of interest.

15 Hens are treated with estradiol to stimulate oviduct maturation as described in Oka and Schimke (T. Oka and RT Schimke, *J. Cell Biol.*, 41, 816 (1969)), Palmiter, Christensen and Schimke (*J. Biol. Chem.* 245(4):833-845, 1970). Specifically, repeated daily injections of 1 mg estradiol benzoate are performed sometime before the onset of sexual maturation, a period ranging from 1 – 14 weeks of age. After a 20 stimulation period sufficient to maximize development of the oviduct, hormone treatment is withdrawn thereby causing regression in oviduct secretory cell size but not cell number. At an optimum time after hormone withdrawal, the oviducts of treated hens are injected with the transposon-based vector. Hens are subjected to additional estrogen stimulation after an optimized time during which the transposon-based vector is taken up into oviduct secretory cells. Re-stimulation by estrogen activates the transposon mechanism of the transposon-based vector, causing the 25 integration of the gene of interest into the host genome. Estrogen stimulation is then withdrawn and hens continue normal sexual development. If a developmentally regulated promoter such as the ovalbumin promoter is used, expression of the 30 transposon-based vector initiates in the oviduct at the time of sexual maturation. Intra-ovarian artery injection during this window allows for high and uniform transfection efficiencies of ovarian follicles to produce germ-line transfections and possibly oviduct expression.

Other means are also used to synchronize the development, or regression, of the oviduct and ovary to allow high and uniform transfection efficiencies. Alterations of lighting and/or feed regimens, for example, cause hens to 'molt' during which time the oviduct and ovary regress. Molting is used to synchronize hens for transfection, 5 and may be used in conjunction with other hormonal methods to control regression and/or development of the oviduct and ovary.

EXAMPLE 17

Isolation of Human Proinsulin Using Anti-TAG Column Chromotography

10 A HiTrap NHS-activated 1 mL column (Amersham) was charged with a 30 amino acid peptide that contained the gp-41 epitope containing gp-41's native disulfide bond that stabilizes the formation of the gp-41 hairpin loop. The 30 amino acid gp41 peptide is provided as SEQ ID NO:23. Approximately 10 mg of the peptide was dissolved in coupling buffer (0.2 M NaHCO₃, 0.5 M NaCl, pH 8.3 and 15 the ligand was circulated on the column for 2 hours at room temperature at 0.5 mL/minute. Excess active groups were then deactivated using 6 column volumes of 0.5 M ethanolamine, 0.5 M NaCl, pH 8.3 and the column was washed alternately with 6 column volumes of acetate buffer (0.1 M acetate, 0.5 M NaCl, pH 4.0) and ethanolamine (above). The column was neutralized using 1 X PBS. The column was 20 then washed with buffers to be used in affinity purification: 75 mM Tris, pH 8.0 and elution buffer, 100 mM glycine-HCl, 0.5 M NaCl, pH 2.7. Finally, the column was equilibrated in 75 mM Tris buffer, pH 8.0.

Antibodies to gp-41 were raised in goats by inoculation with the gp-41 peptide described above. More specifically, goats were inoculated, given a booster injection 25 of the gp-41 peptide and then bled. Serum was harvested by centrifugation. Approximately 30 mL of goat serum was filtered to 0.45 uM and passed over a TAG column at a rate of 0.5 mL/min. The column was washed with 75 mM Tris, pH 8.0 until absorbance at 280 nm reached a baseline. Three column volumes (3 mL) of elution buffer (100 mM glycine, 0.5 M NaCl, pH 2.7) was applied, followed by 75 30 mM Tris buffer, pH 8.0, all at a rate of 0.5 mL/min. One milliliter fractions were collected. Fractions were collected into 200 uL 1 M Tris, pH 9.0 to neutralize acidic fractions as rapidly as possible. A large peak eluted from the column, coincident with the application the elution buffer. Fractions were pooled. Analysis by SDS-PAGE

showed a high molecular weight species that separated into two fragments under reducing condition, in keeping with the heavy and light chain structure of IgG.

Pooled antibody fractions were used to charge two 1 mL HiTrap NHS-activated columns, attached in series. Coupling was carried out in the same manner as

5 that used for charging the TAG column.

Isolation of Ovalbumin-TAG-Proinsulin from Egg White

Egg white from quail and chickens treated by intra-oviduct injection of the CMV-ovalbumin-TAG-proinsulin construct were pooled. Viscosity was lowered by subjecting the allantoid fluid to successively finer pore sizes using negative pressure

10 filtration, finishing with a 0.22 μ M pore size. Through the process, egg white was diluted approximately 1:16. The clarified sample was loaded on the Anti-TAG column and eluted in the same manner as described for the purification of the anti-TAG antibodies. A peak of absorbance at 280 nm, coincident with the application of the elution buffer, indicated that protein had been specifically eluted from the Anti-

15 TAG column. Fractions containing the eluted peak were pooled for analysis.

The pooled fractions from the Anti-TAG affinity column were characterized by SDS-PAGE and western blot analysis. SDS-PAGE of the pooled fractions revealed a 60 kDa molecular weight band not present in control egg white fluid, consistent with the predicted molecular weight of the transgenic protein. Although some

20 contaminating bands were observed, the 60 kDa species was greatly enriched compared to the other proteins. An aliquot of the pooled fractions was cleaved overnight at room temperature with the protease, enterokinase. SDS-PAGE analysis of the cleavage product, revealed a band not present in the uncut material that co-migrated with a commercial human proinsulin positive control. Western blot analysis

25 showed specific binding to the 60 kDa species under non-reducing condition (which preserve the hairpin epitope of gp-41 by retaining the disulfide bond). Western analysis of the low molecular weight species that appeared upon cleavage with an anti-human proinsulin antibody, conclusively identified the cleaved fragment as human proinsulin.

30

EXAMPLE 18

Construction of a Transposon-based Transgene for the Expression of a Monoclonal Antibody

Production of a monoclonal antibody using transposon-based transgenic methodology is accomplished in a variety of ways.

- 5 1) two vectors are constructed: one that encodes the light chain and a second vector that encodes the heavy chain of the monoclonal antibody. These vectors are then incorporated into the genome of the target animal by at least one of two methods: a) direct transfection of a single animal with both vectors (simultaneously or as separate events); or, b) a male and a female of the species carry in their germline one of the vectors and then they are mated to produce progeny that inherit a copy of each.
- 10 2) the light and heavy chains are included on a single DNA construct, either separated by insulators and expression is governed by the same (or different) promoters, or by using a single promoter governing expression of both transgenes with the inclusion of elements that permit separate transcription of both transgenes, such as an internal ribosome entry site.
- 15

The following example describes the production of a transposon-based DNA construct that contains both the coding region for a monoclonal light chain and a heavy chain on a single construct. Beginning with the vector pTnMod, the coding sequences for the heavy and light chains are added, each preceded by an appropriate promoter and signal sequence. Using methods known to one skilled in the art, approximately 1 Kb of the proximal elements of the ovalbumin promoter are linked to the signal sequence of ovalbumin or some other protein secreted from the target tissue. Two copies of the promoter and signal sequence are added to the multiple cloning site of pTnMod, leaving space and key restriction sites between them to allow the subsequent addition of the coding sequences of the light and heavy chains of the monoclonal antibody. Methods known to one skilled in the art allow the coding sequences of the light and heavy chains to be inserted in-frame for appropriate expression. For example, the coding sequence of light and heavy chains of a murine monoclonal antibody that show specificity for human seminoprotein have recently been disclosed (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively). The light chain cDNA sequence is provided in SEQ

ID NO:34, whereas the cDNA of the heavy chain is reported as provided in SEQ ID NO:35.

Thus one skilled in the art can produce both the heavy and light chains of a monoclonal antibody in a single cell within a target tissue and species. If the modified cell contained normal posttranslational modification capabilities, the two chains would form their native configuration and disulfide attachments and be substrates for glycosylation. Upon secretion, then, the monoclonal antibody is accumulated, for example, in the egg white of a chicken egg, if the transgenes are expressed in the magnum of the oviduct.

It should also be noted that, although this example details production of a full-length murine monoclonal antibody, the method is quite capable of producing hybrid antibodies (e.g. a combination of human and murine sequences; 'humanized' monoclonal antibodies), as well as useful antibody fragments, known to one skilled in the art, such as Fab, Fc, F(ab) and Fv fragments. This method can be used to produce molecules containing the specific areas thought to be the antigen recognition sequences of antibodies (complementarity determining regions), linked, modified or incorporated into other proteins as desired.

EXAMPLE 19

Treatment of rats with a transposon-based vector for tissue-specific insulin gene incorporation

Rats are made diabetic by administering the drug streptozotocin (Zanosar; Upjohn, Kalamazoo, MI) at approximately 200 mg/kg. The rats are bred and maintained according to standard procedures. A transposon-based vector containing a proinsulin gene, an appropriate carrier, and, optionally, a transfection agent, are injected into rats' singhepatic (if using G6P) artery with the purpose of stable transformation. Incorporation of the insulin gene into the rat genome and levels of insulin expression are ascertained by a variety of methods known in the art. Blood and tissue samples from live or sacrificed animals are tested. A combination of PCR, Southern and Northern blots, *in-situ* hybridization and related nucleic acid analysis methods are used to determine incorporation of the vector-derived proinsulin DNA and levels of transcription of the corresponding mRNA in various organs and tissues of the rats. A combination of SDS-PAGE gels, Western Blot analysis, radioimmunoassay, and ELISA and other methods known to one of ordinary skill in

the art are used to determine the presence of insulin and the amount produced. Additional transfections of the vector are used to increase protein expression if the initial amounts of the expressed insulin are not satisfactory, or if the level of expression tapers off. The physiological condition of the rats is closely examined 5 post-transfection to register positive or any negative effects of the gene therapy. Animals are examined over extended periods of time post-transfection in order to monitor the stability of gene incorporation and protein expression.

EXAMPLE 20

10

Exemplary Transposon-Based Vectors

The following example provides a description of various transposon-based vectors of the present invention and several constructs for insertion into the transposon-based vectors of the present invention. These examples are not meant to 15 be limiting in any way. The constructs for insertion into a transposon-based vector are provided in a cloning vector labeled pTnMCS.

pTnMCS (base vector)

Bp 1 – 130 Remainder of F1 (-) ori of pBluescriptII sk(-) (Stratagene) bp1-130
20 Bp 133 – 1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems) bp2 29-1873
Bp 1783 – 2991 Transposase, from Tn10 (GenBank accession #J01829) bp 108-1316
Bp 2992 – 3344 Non coding DNA from vector pNK2859
Bp 3345 – 3387 Lambda DNA from pNK2859
25 Bp 3388 – 3457 70 bp of IS10 left from Tn10
Bp 3464 – 3670 Multiple cloning site from pBluescriptII sk(-), thru the XmaI site bp924-718
Bp 3671 - 3715 Multiple cloning site from pBluescriptII sk(-), from the XmaI site thru the XhoI site. These base pairs are usually lost when cloning into pTnMCS.bp
30 717-673
Bp 3716 – 4153 Multiple cloning site from pBluescriptII sk(-), from the XhoI site bp672-235
Bp 4159 - 4228 70 bp of IS10 left from Tn10
Bp 4229 - 4270 Lambda DNA from pNK2859

Bp 4271 – 5114 Non-coding DNA from pNK2859

Bp 5115 - 7315 pBluescript sk (-) base vector (Stratagene, Inc.) bp 761-2961

pTnMCS (CMV-prepro-ent-hGH-CPA)

5 Bp 1 – 3670 from vector PTnMCS, bp 1 - 3670
Bp 3676 – 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864
Bp 5326 - 5496 Capsite/Prepro taken from GenBank accession # X07404, bp 563 – 733
10 Bp 5504 - 5652 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
Bp 5653 – 6306 Human growth hormone taken from GenBank accession # V00519, bp 1-654
Bp 6313 – 6720 Conalbumin polyA taken from GenBank accession # Y00407, bp
15 10651-11058
Bp 6722 – 10321 from cloning vector pTnMCS, bp 3716-7315

pTnMCS (CMV-CHOVg-ent-ProInsulin-synPA) (SEQ ID NO:41)

Bp 1 – 3670 from vector PTnMCS, bp 1 - 3670
20 Bp 3676 – 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864
Bp 5327 –6480 Chicken ovalbumin gene taken from GenBank accession # V00383, bp 66-1219
Bp 6487 - 6636 Synthetic spacer sequence and hairpin loop of HIV gp41 with an
25 added enterokinase cleavage site
Bp 6637 – 6897 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377
Bp 6898 – 6942 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)
30 Bp 6943 - 7295 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271
Bp 7296 – 10895 from cloning vector pTnMCS, bp 3716-7315

pTnMCS (CMV-prepro-ent-ProInsulin-synPA)

Bp 1 - 3670 from vector PTnMCS, bp 1 - 3670

Bp 3676 - 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864

5 Bp 5326 - 5496 Capsite/Prepro taken from GenBank accession # X07404, bp 563 - 733

Bp 5504 - 5652 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

10 Bp 5653 - 5913 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377

Bp 5914 - 5958 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)

Bp 5959 - 6310 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271

15 Bp 6313 - 9912 from cloning vector pTnMCS, bp 3716-7315

pTnMCS(Chicken OVep+OVg'+ENT+proins+syn polyA)

Bp 1 - 3670 from vector pTnMCS, bp 1 - 3670

20 Bp 3676 - 4350 Chicken Ovalbumin enhancer taken from GenBank accession #S82527.1 bp 1-675

Bp 4357 - 5692 Chicken Ovalbumin promoter taken from GenBank accession # J00895M24999 bp 1-1336

25 Bp 5699 - 6917 Chicken Ovalbumin gene from GenBank Accession # V00383.1 bp 2-1220. (This sequence includes the 5'UTR, containing putative cap site, bp 5699-5762.)

Bp 6924 - 7073 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 7074 - 7334 Human proinsulin GenBank Accession # NM000207 bp 117-377

30 Bp 7335 - 7379 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 7380 - 7731 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 7733 - 11332 from vector pTnMCS, bp 3716 - 7315

pTnMCS(Chicken OVep+prepro+ENT+proins+syn polyA)

Bp 1 – 3670 from cloning vector pTnMCS, bp 1 - 3670

Bp 3676 – 4350 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675

Bp 4357 – 5692 Chicken Ovalbumin promoter taken from GenBank accession # J00895-M24999 bp 1-1336

Bp 5699 – 5869 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-733

10 Bp 5876 - 6025 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6026 - 6286 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 6287 - 6331 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

15 Bp 6332 - 6683 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 6685 – 10284 from cloning vector pTnMCS, bp 3716 - 7315

pTnMCS(Quail OVep+OVg'+ENT+proins+syn polyA)

20 Bp 1 – 3670 from cloning vector pTnMCS, bp 1 - 3670

Bp 3676 – 4333 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession # S82527.1, bp 1-675. (There are multiple base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)

Bp 4340 – 5705 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)

30 Bp 5712 – 6910 Quail Ovalbumin gene, EMBL accession # X53964, bp 1-1199. (This sequence includes the 5'UTR, containing putative cap site bp 5712-5764.)

Bp 6917 - 7066 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 7067 - 7327 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 7328 - 7372 Spacer DNA, derived as an artifact from the cloning vectors pTOPO

5 Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 7373 - 7724 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 7726 - 11325 from cloning vector pTnMCS, bp 3716 - 7315

10 pTnMCS (CHOVep-prepro-ent-hGH-CPA)
Bp 1 - 3670 from vector PTnMCS, bp 1-3670
Bp 3676 - 4350 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1, bp 1-675
Bp 4357 - 5692 Chicken Ovalbumin promoter taken from GenBank accession #

15 J00899-M24999, bp 1-1336
Bp 5699 - 5869 Capsite/Prepro taken from GenBank accession # X07404, bp 563-733
Bp 5877 - 6025 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
Bp 6026 - 6679 Human growth hormone taken from GenBank accession # V00519,

20 bp 1-654
Bp 6686 - 7093 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058
Bp 7095 - 10694 from cloning vector pTnMCS, bp 3716-7315

25 pTnMCS(Quail OVep+prepro+ENT+proins+syn polyA)
Bp 1 - 3670 from cloning vector pTnMCS, bp 1 - 3670
Bp 3676 - 4333 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far- upstream chicken ovalbumin enhancer, GenBank accession #S82527.1, bp 1-675. (There are multiple base pair

30 substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)
Bp 4340 - 5705 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter,

GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)

Bp 5712 – 5882 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-

5 733

Bp 5889 - 6038 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6039 - 6299 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 6300 - 6344 Spacer DNA, derived as an artifact from the cloning vectors pTOPO

10 Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 6345 - 6696 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 6698 – 10297 from cloning vector pTnMCS, bp 3716 - 7315

15 PTnMOD

Bp 1 - 130 remainder of F1 (-) ori of pBluescriptII sk(-) (Stratagene) bp1-130

Bp 133 - 1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems) bp229-1873

Bp 1783 - 2991 Transposase, modified from Tn10 (GenBank accession #J01829) bp

20 108-1316

Bp 2992 - 2994 Engineered stop codon

Bp 2996 - 3411 Synthetic polyA from gWIZ (Gene Therapy Systems) bp 1922 - 2337

Bp 3412 - 3719 Non-coding DNA from vector pNK2859

Bp 3720 - 3762 Lambda DNA from pNK2859

25 Bp 3763 - 3832 70 bp of IS10 left from Tn10

Bp 3839 – 4045 Multiple cloning site from pBluescriptII sk(-), thru the XmaI site bp 924-718

Bp 4046 - 4090 Multiple cloning site from pBluescriptII sk(-), from the XmaI site thru the XhoI site. These base pairs are usually lost when cloning into pTnMCS. bp

30 717-673

Bp 4091 – 4528 Multiple cloning site from pBluescriptII sk(-), from the XhoI site bp 672-235

Bp 4534 - 4603 70 bp of IS10 left from Tn10
Bp 4604 - 4645 Lambda DNA from pNK2859
Bp 4646 - 5489 Non-coding DNA from pNK2859
Bp 5490 - 7690 pBluescript sk (-) base vector (Stratagene, INC) bp 761-2961

5

pTnMOD (CHOVep-prepro-ent-hGH-CPA)

Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045
Bp 4051 - 4725 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1, bp 1 - 675
10 Bp 4732 - 6067 Chicken Ovalbumin promoter taken from GenBank accession # J00899-M24999, bp 1-1336
Bp 6074 - 6245 Capsite/Prepro taken from GenBank accession # X07404, bp 563 - 733
Bp 6252 - 6400 Synthetic spacer sequence and hairpin loop of HIV gp41 with an
15 added enterokinase cleavage site
Bp 6401 - 7054 Human growth hormone taken from GenBank accession # V00519, bp 1-654
Bp 7061 - 7468 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058
20 Bp 7470 - 11069 from cloning vector pTnMCS, bp 3716-7315

pTnMOD (CMV-CHOVg-ent-ProInsulin-synPA) (SEQ ID NO:42)

Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045
Bp 4051 - 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy
25 systems), bp 230-1864
Bp 5702 - 6855 Chicken ovalbumin gene taken from GenBank accession # V00383, bp 66-1219
Bp 6862 - 7011 Synthetic spacer sequence and hairpin loop of HIV gp41 with an
added enterokinase cleavage site
30 Bp 7012 - 7272 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377
Bp 7273 - 7317 Spacer DNA, derived as an artifact from the cloning vectors pTOPO
Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)

Bp 7318 - 7670 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271

Bp 7672 - 11271 from cloning vector pTnMCS, bp 3716-7315

5 pTnMOD (CMV-prepro-ent-hGH-CPA)

Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045

Bp 4051 - 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230-1864

Bp 5701 - 5871 Capsite/Prepro taken from GenBank accession # X07404, bp 563 -

10 733

Bp 5879 - 6027 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6028 - 6681 Human growth hormone taken from GenBank accession # V00519, bp 1-654

15 Bp 6688 - 7095 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058

Bp 7097 - 10696 from cloning vector pTnMCS, bp 3716-7315

pTnMOD (CMV-prepro-ent-ProInsulin-synPA)

20 Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045

Bp 4051 - 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230-1864

Bp 5701 - 5871 Capsite/Prepro taken from GenBank accession # X07404, bp 563 - 733

25 Bp 5879 - 6027 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6028 - 6288 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377

Bp 6289 - 6333 Spacer DNA, derived as an artifact from the cloning vectors pTOPO

30 Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)

Bp 6334 - 6685 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271

Bp 6687 - 10286 from cloning vector pTnMCS, bp 3716-7315

pTnMOD(Chicken OVep+OVg'+ENT+proins+syn polyA) (SEQ ID NO:43)

Bp 1 – 4045 from cloning vector pTnMOD, bp 1 - 4045

Bp 4051 – 4725 Chicken Ovalbumin enhancer taken from GenBank accession #

5 S82527.1 bp 1-675

Bp 4732 – 6067 Chicken Ovalbumin promoter taken from GenBank accession #

J00895-M24999 bp 1-1336

Bp 6074 – 7292 Chicken Ovalbumin gene from GenBank Accession # V00383.1 bp

2-1220. (This sequence includes the 5'UTR, containing putative cap site bp 6074-

10 6137.)

Bp 7299 - 7448 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 7449 - 7709 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 7710 - 7754 Spacer DNA, derived as an artifact from the cloning vectors pTOPO

15 Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 7755 - 8106 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 8108 – 11707 from cloning vector pTnMCS, bp 3716 - 7315

20 pTnMOD(Chicken OVep+prepro+ENT+proins+syn polyA)

Bp 1 – 4045 from cloning vector pTnMCS, bp 1 - 4045

Bp 4051 – 4725 Chicken Ovalbumin enhancer taken from GenBank accession #

S82527.1 bp 1-675

Bp 4732 – 6067 Chicken Ovalbumin promoter taken from GenBank accession #

25 J00895-M24999 bp 1-1336

Bp 6074 – 6244 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-733

Bp 6251 - 6400 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

30 Bp 6401 - 6661 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 6662 - 6706 Spacer DNA, derived as an artifact from the cloning vectors pTOPO

Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 6707 - 7058 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 7060 - 10659 from cloning vector pTnMCS, bp 3716 - 7315

5 pTnMOD(Quail OVEp+OVg'+ENT+proins+syn polyA)

Bp 1 - 4045 from cloning vector pTnMCS, bp 1 - 4045

Bp 4051 - 4708 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession # S82527.1, bp 1-675. (There are multiple base pair

10 substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)

Bp 4715 - 6080 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair

15 substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)

Bp 6087 - 7285 Quail Ovalbumin gene, EMBL accession # X53964, bp 1-1199. (This sequence includes the 5'UTR, containing putative cap site bp 6087-6139.)

Bp 7292 - 7441 Synthetic spacer sequence and hairpin loop of HIV gp41 with an 20 added enterokinase cleavage site

Bp 7442 - 7702 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 7703 - 7747 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 7748 - 8099 Synthetic polyA from the cloning vector gWIZ (Gene Therapy 25 Systems) bp 1920 - 2271

Bp 8101 - 11700 from cloning vector pTnMCS, bp 3716 - 7315

pTnMOD(Quail OVEp+prepro+ENT+proins+syn polyA)

Bp 1 - 4045 from cloning vector pTnMCS, bp 1 - 4045

30 Bp 4051 - 4708 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession # S82527.1, bp 1-675. (There are multiple

base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)

Bp 4715 – 6080 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter,

5 GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)

Bp 6087 – 6257 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-733

10 Bp 6264 - 6413 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6414 - 6674 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 6675 - 6719 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

15 Bp 6720 - 7071 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 7073 – 10672 from cloning vector pTnMCS, bp 3716 - 7315

PTnMod(CMV/Transposase/ChickOvep/prepro/ProteinA/ConpolyA)

20 BP 1-130 remainder of F1 (-) ori of pBluescriptII sk(-) (Stragagene) bp 1-130.

BP 133-1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems) bp 229-1873.

BP 1780-2987 Transposase, modified from Tn10 (GenBank #J01829).

BP 2988-2990 Engineered stop codon.

25 BP 2991-3343 non coding DNA from vector pNK2859.

BP 3344-3386 Lambda DNA from pNK2859.

BP 3387-3456 70bp of IS10 left from Tn10.

BP 3457-3674 multiple cloning site from pBluescriptII sk(-) bp 924-707.

BP 3675-5691 Chicken Ovalbumin enhancer plus promoter from a Topo Clone 10

30 maxi 040303 (5' XmaI, 3' BamHI)

BP 5698-5865 prepro with Cap site amplified from cecropin of pMON200 GenBank # X07404 (5'BamHI, 3'KpnI)

BP 5872-7338 Protein A gene from GenBank# J01786, mature peptide bp 292-1755
(5'KpnI, 3'SacII)

BP 7345-7752 ConPolyA from Chicken conalbumin polyA from GenBank # Y00407
bp 10651-11058. (5'SacII, 3'Xhol)

5 BP 7753-8195 multiple cloning site from pBluescriptII sk(-) bp 677-235.

BP 8196-8265 70 bp of IS10 left from Tn10.

BP 8266-8307 Lamda DNA from pNK2859

BP 8308-9151 noncoding DNA from pNK2859

BP 9152-11352 pBluescriptII sk(-) base vector (Stratagene, INC.) bp 761-2961

10

All patents, publications and abstracts cited above are incorporated herein by reference in their entirety. It should be understood that the foregoing relates only to preferred embodiments of the present invention and that numerous modifications or alterations may be made therein without departing from the spirit and the scope of the

15 present invention as defined in the following claims.

CLAIMS

We Claim:

1. A vector comprising:
 - 5 a) a transposase gene operably linked to a first promoter; and
 - b) one or more genes of interest operably-linked to one or more additional promoters;
 - c) wherein the one or more genes of interest and their operably-linked promoters are flanked by transposase insertion sequences recognized by the transposase; and
 - 10 d) wherein the first promoter comprises a modified Kozak sequence comprising ACCATG.
2. The vector of claim 1, wherein one to twenty codons at a beginning of
15 the transposase gene are modified by changing a nucleotide at a third base position of the codon to an adenine or thymine without modifying the amino acid encoded by the codon.
3. The vector of claim 2, wherein the transposase gene is modified in its
20 first ten codons.
4. The vector of claim 1, wherein the transposase is a *Tn10* transposase.
5. The vector of claim 1, wherein the first promoter is a constitutive
25 promoter.
6. The vector of claim 1, wherein the first promoter is an inducible promoter.
- 30 7. The vector of claim 6, wherein the inducible promoter is an ovalbumin promoter or a vitellogenin promoter.
8. The vector of claim 1, wherein one gene of interest is operably-linked to a second promoter.
- 35 9. The vector of claim 8, wherein the second promoter is a constitutive promoter.

10. The vector of claim 8, wherein the second promoter is an inducible promoter.
- 5 11. The vector of claim 10, wherein the inducible promoter is an ovalbumin promoter or a vitellogenin promoter.
12. The vector of claim 1, further comprising a polyA sequence operably-linked to the transposase gene.
- 10 13. The vector of claim 12, wherein the polyA sequence is a conalbumin polyA sequence.
14. The vector of claim 1 or claim 12, further comprising two stop codons operably-linked to the transposase gene.
- 15 15. The vector of claim 1, wherein a first gene of interest is operably-linked to a second promoter and a second gene of interest is operably-linked to a third promoter.
- 20 16. The vector of claim 1, wherein a first and a second gene of interest are operably-linked to a second promoter.
17. The vector of claim 1, further comprising an enhancer operably-linked to the one or more genes of interest.
- 25 18. The vector of claim 17, wherein the enhancer comprises at least a portion of an ovalbumin enhancer.
19. The vector of claim 1, further comprising an egg directing sequence operably-linked to the one or more genes of interest.
- 30 20. The vector of claim 19, wherein the egg directing sequence is an ovalbumin signal sequence or an ovomucoid signal sequence.
- 35 21. The vector of claim 19, wherein the egg directing sequence is a vitellogenin targeting sequence.

22. A method for producing a transgenic animal comprising administering to the animal a vector comprising:

- a) a modified transposase gene operably linked to a first promoter; and
- 5 b) one or more genes of interest operably-linked to one or more additional promoters;
- c) wherein the one or more genes of interest and their operably-linked promoters are flanked by transposase insertion sequences recognized by the transposase; and
- 10 d) wherein the first promoter comprises a modified Kozak sequence comprising ACCATG.

23. The method of claim 22, wherein one to twenty codons at a beginning of the transposase gene are modified by changing a nucleotide at a third base position of the codon to an adenine or thymine without modifying the amino acid encoded by the codon.

15

24. The method of claim 23, wherein the transposase gene is modified in its first ten codons.

20

25. The method of claim 22, wherein the vector is administered via an intratesticular, intraarterial, intraoviductal or intraembryonic route.

26. The method of claim 22, wherein the transposase is a Tn10 transposase.

25

27. The method of claim 22, wherein the first promoter is a constitutive promoter.

30

28. The method of claim 22, wherein the first promoter is an inducible promoter.

35

29. The method of claim 28, wherein the inducible promoter is selected from the group consisting of an ovalbumin promoter, an ovomucoid promoter and a vitellogenin promoter.

30. The method of claim 22, wherein one gene of interest is operably linked to a second promoter.

31. The method of claim 30, wherein the second promoter is an inducible promoter.

5 32. The method of claim 31, wherein the inducible promoter is an ovalbumin promoter or a vitellogenin promoter.

33. The method of claim 22, wherein the vector further comprises a polyA sequence operably-linked to the transposase gene.

10 34. The method of claim 33, wherein the polyA sequence is a conalbumin polyA sequence.

35. The method of claim 22 or claim 33, wherein the vector further comprises two stop codons operably-linked to the transposase gene.

15 36. The method of claim 22, wherein a first gene of interest is operably-linked to a second promoter and a second gene of interest is operably-linked to a third promoter.

20 37. The method of claim 22, wherein a first and second gene of interest are operably linked to a second promoter.

38. The method of claim 22, further comprising an enhancer operably-linked to the one or more genes of interest.

25 39. The method of claim 38, wherein the enhancer comprises at least a portion of an ovalbumin enhancer.

40. The method of claim 22, wherein the animal is an avian animal.

30 41. The method of claim 40, wherein the avian animal is a chicken or a quail.

35 42. An egg produced by the avian animal of claim 40, wherein the egg contains one or more desired proteins encoded by the one or more genes of interest.

43. A transgenic sperm produced by the animal of claim 22.

44. A method of producing a desired protein comprising:

- 5 a) administering to an animal a vector comprising a modified transposase gene operably linked to a first promoter, and a gene of interest encoding the desired protein operably-linked to a second promoter; and
- b) isolating the desired protein produced in the animal; wherein
- c) the gene of interest and its operably-linked promoter are flanked by transposase insertion sequences recognized by the transposase; and
- 10 d) the first promoter comprises a modified Kozak sequence comprising ACCATG.

45. The method of claim 44, wherein one to twenty codons at a beginning 15 of the transposase gene are modified by changing a nucleotide at a third base position of the codon to an adenine or thymine without modifying the amino acid encoded by the codon.

46. The method of claim 44, wherein the animal is an egg-laying animal 20 and the desired protein is isolated from an egg white.

47. The method of claim 44, wherein the vector further comprises a TAG 25 sequence and wherein the desired protein is purified using the TAG sequence.

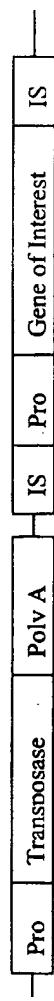
48. The method of claim 47, wherein the TAG sequence comprises a 30 polynucleotide sequence encoding an antigenic portion of a gp41 protein, an enterokinase cleavage site and a spacer polynucleotide sequence.

49. The method of claim 47, wherein the TAG sequence comprises a 35 polynucleotide sequence shown in SEQ ID NO:22.

50. The method of claim 44, wherein the desired protein is a lytic protein.

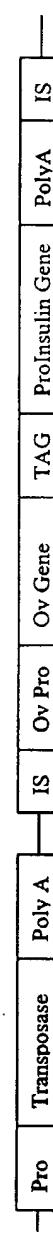
51. The method of claim 44, wherein the vector further comprises a second 35 gene of interest operably-linked to a third promoter and wherein the genes of interest encode antibody polypeptides.

FIGURE 1



2/7

FIGURE 2



3/7

FIGURE 3



FIGURE 4



5/7

FIGURE 5

IS	Oval Pro	prepro	Heavy chain	pro	Light chain	polyA	IS
----	----------	--------	-------------	-----	-------------	-------	----

6/7

FIGURE 6

IS	Oval Pro	prepro	Light chain	ent	Heavy chain	polyA	IS
----	----------	--------	-------------	-----	-------------	-------	----

FIGURE 7

A.

Tail-to-Tail

IS	Oval Pro	Oval SS	Light chain	Poly A	Spacer DNA	Poly A	Heavy chain	Oval SS	Oval Pro	IS
----	----------	---------	-------------	--------	------------	--------	-------------	---------	----------	----

B.

Tail-to-Head

IS	Oval Pro	Oval SS	Light chain	Poly A	Spacer DNA	Poly A	Heavy chain	Oval SS	Poly A	IS
----	----------	---------	-------------	--------	------------	--------	-------------	---------	--------	----

Appendix A

SEQ ID NO:1 (pTnMod)

5 CTGACGCCCTGTAGCCGC GCATTAAGCG CGCGGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCCTTCGC 100
 TTTCTTCCTCCTCAGCAGTCAGTACATCATAA TATGTACATT TATATTGGCT 150
 GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
 CATGTCACAC ATTACGCCA TGTTGACATT GATTATTGAC TAGTTATTAA 250
 10 TAGTAATCAA TTACGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG 300
 CGTTACATCAA CTTACGGTAA ATGGCCCGCC TGCGACGGC CCCAACGACC 350
 CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCAATA 400
 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AACTGCCA 450
 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG 500
 15 TCAATGACGG TAAATGGCC GCCTGGCATT ATGCCAGTA CATGACCTTA 550
 TGGGACTTTCTACTTGCA GTACATCTAC GTATTAGTC TCGCTATTAC 600
 CATGGTGATC CGGTTTTGGC ACTACATCAA TGCGCTGGA TAGCGGTTTG 650
 ACTCACGGGG ATTTCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG 700
 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC 750
 20 CCCATTGACG CAAATGGCG GTAGGGTGT ACGGTGGAG GTCTATATAA 800
 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACCG 850
 TGTTTTGACC TCCATAGAAG ACACGGGAC CGATCAGGCC TCCGCGGCCG 900
 GGAACGGTGC ATTGGAACGC GGATTCCCG TGCCAAGAGT GACGTAAGTA 950
 CGGCCTATAG ACTCTATAGG CACACCCCT TGCTCTTAT GCATGCTATA 1000
 25 CTGTTTTGG CTTGGGGCCT ATACACCCCCC GCTTCCTTAT GCTATAGGTG 1050
 ATGGTATAGC TTAGCCTATA GGTGTGGTT ATTGACCATT ATTGACCACT 1100
 CCCCTATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
 GCCACAACTA TCTCTATGG CTATATGCCA ATACTCTGTC CTTCAGAGAC 1200
 TGACACGGAC TCTGTATTT TACAGGATGG GGTCCCATTT ATTATTACAA 1250
 30 AATTCAACATA TACAACAACG CCGTCCCCCG TGCGCCAGT TTTTATTAAA 1300
 CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTT CGGACATGGG 1350
 CTCTTCTCCG GTAGCGGGCG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
 CTCCAGCGGC TCATGGTCGC TCGGCAGCTC CTTGCTCCTA ACAGTGGAGG 1450
 CCAGACTTAG GCACAGCACA ATGCCACCA CCACCAAGTGT GCCGCACAAG 1500
 GCGCGTGGCG TAGGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGGCTCG 1550
 CACCGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG 1600
 GCAGCTGAGT TGTGTATTC TGATAAGAGT CAGAGGTAAC TCCCCTTGCG 1650
 GTGCTGTAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TGTTGCTGC 1700
 CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTTT 1750
 40 CCATGGGTCT TTTCTGCACT CACCGTCGGA CCATGTGTCG ACTTGATATT 1800
 TTACATGATT CTCTTACCA ATTCTGCCCG GAATTACACT TAAAACGACT 1850
 CAACAGCTTA ACGTTGGCTT GCCACGCATT ACTTGACTGT AAAACTCTCA 1900
 CTCTTACCGA ACTTGGCCGT AACCTGCCA CCAAAGCGAG AACAAAACAT 1950
 AACATCAACG GAATCGACCG ATTGTAGGT AATCGTCACC TCCACAAAGA 2000
 45 GCGACTCGCT GTATACCGTT GGCATGCTAG CTTATCTGT TCGGGAATAC 2050
 GATGCCATT GTACTGGTT ACTGGCTCTGA TATCGTGTGAG CAAAACGAC 2100
 TTATGGTATT GCGAGCTTCA GTCCGACTAC ACAGTCGTT TCCTACTCTT 2150
 TATGAGAAAG CGTCTCCGCT TTCAAGTAA TGTTCAAGA AAGCTCATGA 2200
 CCAATTCTA GCGGACCTTG CGAGCATTCT ACCAGTAAC ACCACACCGC 2250
 50 TCATTGTCAG TGATGCTGGC TTTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
 AAGCTGGTT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
 AGACCTAGGA GCGGAAAAGT GGAAACCTAT CAGCAACTTA CATGATATGT 2400
 CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAAGCAAT 2450
 CCAATCTCAT GCGAAATCT ATTGTATAAA TCTCGCTCTA AAGGCCGAA 2500
 55 AAATCAGCGC TCGACACGGC CTCATTGTC CCACCCGTCA CCTAAATCT 2550
 ACTCAGCGTC GGCAAAAGGAG CCATGGGTTG TAGCAACTAA CTTACCTGTT 2600
 GAAATTGCAA CACCCAAACA ACTTGGTTAT ATCTATTGCA AGCGAATGCA 2650
 GATTGAAGAA ACCTTCCGAG ACTTGGAAAAG TCTTGCTAC GGACTAGGGC 2700
 TACGCCATAG CGAACGAGC AGCTCAGAGC GTTTGTATAT CATGCTGCTA 2750
 60 ATCGCCCTGA TGCTTCACT AACATGTTGG CTTGCGGGCG TTCATGCTCA 2800
 GAAACAAGGT TGGGACAAGC ACTTCCAGGC TAACACAGTC AGAAATCGAA 2850

ACGTACTCTC AACAGTTCGC TTAGGCATGG AAGTTTGC GCGCATTCTGGC 2900
 TACACAATAA CAAGGGAAAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTTATTC ACACATGGTT ACGCTTGGG GAAATTATGA TAATGATCCA 3000
 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 5 TTTTTGTGG ATCTGCTGTG CTTCTAGTT GCCAGCCATC TGTTGTTGC 3100
 CCCTCCCCG TGCCCTCCCT GACCCTGGAA GGTGCCACTC CCACTGTCTC 3150
 TTCTAATAA AATGAGGAAA TTGCATCGCA TTGCTGAGT AGGTGTCTT 3200
 CTATTCTGGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGAA GGATTGGGAA 3250
 10 GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGTGTACCGG TGCTGAAGAA 3400
 TTGACCGGT GACCAAAGGT GCCTTTATC ATCACTTTAA AAATAAAAAAA 3450
 CAATTACTCA GTGCCTGTTA TAAGCAGCAA TTAATTATGA TTGATGCCTA 3500
 CATCACAACA AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 15 TATTTGAACA TTATCTTGAT TATATTATTG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTGGT TTGAATGAAC TTGAAAAAAA 3650
 TTAGCCTGAA ATACATTACT GGTAAGGTAA ACCCCATTGT CAGCAAATTG 3700
 ATCCAAGAGA ACCAACTTAA AGCTTCTCG AGCGGAATGTT AATTCTCGTT 3750
 GACCCCTGAGC ACTGATGAAT CCCCTAATGA TTTGGTAAA AATCATTAAAG 3800
 20 TTAAGGTGGA TACACATCTT GTCATATGAT CCCCTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCCGAG CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGCGC AATTAAACCT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCGCCTCTAG AACTAGTGGA TCCCCCGGGC 4050
 25 TGCAGGAATT CGATATCAAG CTTATCGATA CCGCTGACCT CGAGGGGGGG 4100
 CCCGGTACCC AATTGCGCTT ATAGTGAGTC GTATTACCGG CGCTCACTGG 4150
 CCGTCGTTT ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT 4200
 ATTCGCTTG CAGCACATCC CCCCCTCGCC AGCTGGCGTA ATAGCGAAGA 4250
 GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT 4300
 30 GGAAATTGTA AGCGTTAATA TTTTGTAAA ATTGCGTTA AATTGTTGTT 4350
 AAATCAGCTC ATTTTTAAC CAATAGGCGC AAATCGGCAA AATCCCTTAT 4400
 AAATCAAAG AATAGACCGA GATAGGGTTG AGTGTGTTG CAGTTGGAA 4450
 CAAGAGTCCA CTATTAAAGA ACGTGGACTC CAACTGCAA GGGCGAAAAAA 4500
 CCGTCTATCA GGGCGATGGC CCAACTACTCC GGGATCATAT GACAAGATGT 4550
 35 GTATCACCT TAACCTTAAAT ATTTTTACCA AAATCATTAG GGGATTTCATC 4600
 AGTGCTCAGG GTCAACGAGA ATTAACATTC CGTCAGGAAA GCTTATGATG 4650
 ATGATGTGCT TAAACCTTAACTC CTCAATGGCT GGTTATGATC ATCGCAATAC 4700
 ATGCGAAAAA CCTAAAAGAG CTTGCGATA AAAAGGCCA ATTTATTGCT 4750
 ATTTACCGCG GCTTTTATT GAGCTGAAA GATAAATAAA ATAGATAGGT 4800
 40 TTTAATTGAA GCTAAATCTT CTTTATCGTA AAAATGCC TCTTGGGTTA 4850
 TCAAGAGGGT CATTATATT CGCGGAATAA CATCATTTGG TGACGAAATA 4900
 ACTAAGCACT TGTCTCTGT TTACTCCCT GAGCTTGAGG GGTTAACATG 4950
 AAGGTATCG ATAGCAGGAT AATAATACAG TAAACGCTA AACCAATAAT 5000
 CCAAATCCAG CCATCCAAA TTGGTAGTGA ATGATTATAA ATAACAGCAA 5050
 45 ACAGTAATGG GCAAAATAAC CCGGTGCTG TGGTAAGGCT CACCAATAAT 5100
 CCCTGTAAAG CACCTGTCTG ATGACTCTT GTTGGATAG ACATCACTCC 5150
 CTGTAATGCA GGTAAAGCG TCCCACCAAGGCAATAAA ATTAAAACAG 5200
 GGAAAACCTAA CCAACCTCA GATATAAACG CTTAAAGGC AAATGCACTA 5250
 CTATCTGCAA TAAATCCGAG CAGTACTGCC GTTTTTCGC CCATTTAGTG 5300
 50 GCTATTCTTC CTGCCACAAA GGCTTGGAT ACTGAGTGTAAAGACCAAG 5350
 ACCCGTAATG AAAAGCCAAAC CATCATGCTA TTCAATCATCA CGATTTCTGT 5400
 AATAGCACCA CACCGTGCTG GATTGGCTAT CAATGGCTG AAATAATAAT 5450
 CAACAAATGG CATCGTTAA TAAGTGATGT ATACCGATCA GCTTTGTTG 5500
 55 CTTCTGTGTG AAATTGTTATG CCGCTCACAA TTCCACACAA CATACGAGCC 5600
 GGAAGCATAA AGTGTAAAGC CTGGGGTGCC TAAATGAGTGA GCTAACTCAC 5650
 ATTAATTGCGG TTGCGCTCAC TGCCCGCTT CCAGCTGGGA AACCTGTGCGT 5700
 GCCAGCTGCA TTAATGAATC GGCCAAACGCG CGGGGAGAGG CGGTTGCGT 5750
 ATTGGGCCTC CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTGCGT 5800
 60 TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGTA ATACGGTTAT 5850
 CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG 5900

CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCATAG 5950
 GCTCCGCCCG CCTGACGAGC ATCACAAAAA TCGACGCTCA AGTCAGAGGT 6000
 GGCAGAAACCC GACAGGACTA TAAAGATACC AGGCCTTCCG CCCTGGAAGC 6050
 5 TCCCTCGTGC GCTCTCCTGT TCCGACCCCTG CCGCTTACCG GATACTGTG 6100
 CGCCTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA 6150
 GGTATCTCAG TTCCGGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC 6200
 GAACCCCCCG TTCAGCCCGA CCGCTCGGCC TTATCCGTA ACTATCGTCT 6250
 TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG 6300
 GTAACAGGAT TAGCAGAGCG AGGTATGTTAG GCGGTGCTAC AGAGTTCTTG 6350
 10 AAGTGGTGGC CTAACATACCG CTACACTAGA AGGACAGTAT TTGGTATCTG 6400
 CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT 6450
 CCGGCAAAACA AACCAACCGT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG 6500
 CAGATTACGC GCAGAAAAAA AGGATCTAA GAAGATCCTT TGATCTTTTC 6550
 TACGGGGTCT GACGCTCAGT GGAACGAAAAA CTCACGTTAA GGGATTTGG 6600
 15 TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTT AAATTAACAA 6650
 TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG 6700
 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTTC 6750
 GTTCATCCAT AGTTGCCTGA CTCCCCGTCG TGAGATAAC TACGATACGG 6800
 GAGGGCTTAC CATCTGGCCC CAGTGTGCA ATGATACCGC GAGACCCACG 6850
 20 CTCACCGGCT CCAGATTAT CAGCAATAAA CCAGCCAGGC GGAAGGGCCG 6900
 AGCGCAGAAG TTGGCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAA 6950
 TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA 7000
 CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTGGTA 7050
 TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC 7100
 25 CCCATGTTGT GCAAAAGC GGTTAGCTCC TTGGTCTCTC CGATCGTTGT 7150
 CAGAAGTAAG TTGGCCGCAAG TGTTATCACT CATGGTTATG GCAGCACTG 7200
 ATAATTCTCT TACTGTCATG CCATCCGTA GATGCTTTTC TGTGACTGGT 7250
 GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG 7300
 CTCTGGCCCG CGCTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT 7350
 30 TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGCGAAA ACTCTCAAGG 7400
 ATCTTACCGC TGTTGAGATC CAGTTGATG TAACCCACTC GTGCACCCAA 7450
 CTGATCTTCA GCATCTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAA 7500
 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT 7550
 TGAATACTCA TACTCTTCTT TTTTCAATAT TATTGAAGCA TTTATCAGGG 7600
 35 TTATTGTCTC ATGAGCGGAT ACATATTGA ATGTATTAG AAAATAAAC 7650
 AAATAGGGGT TCCGCGCACA TTCCCCGAA AAGTGCAC 7689

40 SEQ ID NO:2 (PTnMod (CMV/Red))
 CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGGGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCCTTCGC 100
 TTTCTTCCCT TCCCTCTCG CCACGTTCCG CGGCATCAGA TTGGCTATTG 150
 GCCATTGCAAT ACGTTGATC CATATCATAA TATGTACATT TATATTGGCT 200
 45 CATGTCCAACT ATTACCGCA TGTTGACATT GATTATTGAC TAGTTATTAA 250
 TAGTAATCAA TTACGGGTC ATTAGTCAT AGCCCCATATA TGGAGTTCCG 300
 CGTTACATAA CTTACGGTAA ATGGCCGCC TGGCTGACCG CCCAACGACC 350
 CCCGCCCAATT GACGTCATAA ATGACGTATG TTCCCATAGT AACGCCAATA 400
 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTCAGGT AAACGTGCCA 450
 50 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG 500
 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAAGTA CATGACCTTA 550
 TGGGACTTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
 CATGGTGTAT CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG 650
 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG 700
 55 TTTGGCACC AAAATCAACG GGACTTCCA AAATGTCGTA ACAACTCCGC 750
 CCCATTGACG CAAATGGCCG GTAGGGCTGT ACGGTGGGAG GTCTATATAA 800
 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
 TGTGTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG 900
 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA 950
 60 CGGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA 1000
 CTGTTTTGG CTTGGGGCCT ATACACCCCC GCTTCCTTAT GCTATAGGTG 1050

	ATGGTATAGC	TTAGCCTATA	GGTGTGGTT	ATTGACCATT	ATTGACCACT	1100
	CCCTATTGG	TGACGATACT	TTCCATTACT	AATCCATAAC	ATGGCTCTT	1150
	GCCACAACTA	TCTCTATTGG	CTATATGCCA	ATACTCTGTC	CTTCAGAGAC	1200
5	TGACACGGAC	TCTGTATTT	TACAGGATGG	GGTCCCATT	ATTATTACCA	1250
	AATTACACATA	TACAACAACG	CCGTCCCCG	TGCCCGCAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAACATCGG	GTACGTGTT	CGGACATGGG	1350
	CTCTTCTCCG	CTAGCGGCGG	AGCTTCCACA	TCCGAGCCCT	GGTCCCATGC	1400
	CTCCAGCGGC	TCATGGTCGC	TCGGCAGCTC	CTTGCTCCTA	ACAGTGGAGG	1450
10	CCAGACTTAG	GCACAGCACAA	ATGCCACCA	CCACCAAGTGT	GCCGCACAAAG	1500
	GCCGTGGCGG	TAGGGTATGT	GTCTGAAAAT	GAGCGTGGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGGAA	GACTTAAGGC	AGCGGCAGAA	GAAGATGCAG	1600
	GCAGCTGAGT	TGTTGTATTC	TGATAAAGGT	CAGAGGTAAC	TCCCGTTGCG	1650
	GTGCTGTTAA	CGGTGGAGGG	CAGTGTAGTC	TGAGCAGTAC	TCGTTGCTGC	1700
	CGCGCGCGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCTTT	1750
15	CCATGGGTCT	TTTCTGCAGT	CACCGTCGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGATT	CTCTTTACCA	ATTCTGCC	GAATTACACT	TAACAAACGACT	1850
	CAACACCTTA	ACGTTGGCTT	GCCACCCATT	ACTTGACTGT	AAAACCTCTCA	1900
	CTCTTACCGA	ACTTGGCCGT	AACCTGCCAA	CCAAAGCGAG	AACAAAACAT	1950
	AACATCAAAC	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAAGA	2000
20	GCGACTCGCT	GTATACCGTT	GGCATGCTAG	CTTATCTGT	TCGGGAATAC	2050
	GATGCCATT	GTACTTGTG	ACTGGTCTGA	TATTCTGTAG	CAAAAACGAC	2100
	TTATGGTATT	GCGAGCTTCA	GTCGCACTAC	ACGGTCGTT	TGTTACTCTT	2150
	TATGAGAAAG	CGTTCCCGCT	TTCAGAGCA	TGTTCAAAGA	AAGCTCATGA	2200
	CCAATTCTCA	GCCGACCTTG	CGAGCATTCT	ACCGAGTAAC	ACCACACCAC	2250
25	TCATTGTCAG	TGATGCTGGC	TTTAAAGTGC	CATGGTATAA	ATCCGTTGAG	2300
	AAGCTGGGT	GGTACTGGTT	AAGTCGACTA	AGAGGAAAAG	TACAATATGC	2350
	AGACCTAGGA	CGGGAAAAC	GGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
	CATCTAGTCA	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAAGCAAT	2450
	CCAATCTCAT	GCCAAATTCT	ATTGTTAAA	TCTCGCTCTA	AAGGCCGAAA	2500
30	AAATCAGCGC	TCGACACCGA	CTCATGGTCA	CCACCCGTC	CCTAAAATCT	2550
	ACTCAGCGTC	GGCAAAGGGAG	CCATGGGTT	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTGCAA	CACCCAAACCA	ACTTGTAAAT	ATCTATTGCA	AGCGAATGCA	2650
	GATTGAAGAA	ACCTTCCGAG	ACTTGAAAAG	TCCTGCCTAC	GGACTAGGCC	2700
	TACGCCATAG	CGGAACGAGC	AGCTCACAGC	GTGTTGATAT	CATGCTGCTA	2750
35	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	CTTGCGGGCG	TTCATGCTCA	2800
	GAAACAAGGT	TGGGACAAGC	ACTTCCAGGC	TAACACAGTC	AGAAATCGAA	2850
	ACGTACTCTC	AACAGTTCGC	TTAGGCATGG	AAAGTTTGC	GCATTCTGGC	2900
	TACACAATAA	CAAGGGAAAGA	CTTACTCGTG	GCTGCAACCC	TACTAGCTCA	2950
	AAATTTATTC	ACACATGGTT	ACGTTTGGG	GAAATTATGA	TAATGATCCA	3000
40	GATCACTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGTC	TGTGTGTTGG	3050
	TTTTTGTGG	ATCTGCTGTG	CCTCTAGTT	GCCAGCCATC	TGTTGTTGC	3100
	CCCTCCCCCG	TGCTTCCCTT	GACCCCTGAA	GGTGCCACTC	CCACTGTCT	3150
	TTCTCTATAA	AATGAGGAAA	TTGCATCGCA	TTGCTGTAGT	AGGTGTCTATT	3200
	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACAA	GCAAGGGGGA	GGATGGGGAA	3250
45	GACAATAGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAAGG	TGCTGAAGAA	3400
	TTGACCCGGT	GACCAAAGGT	GCCTTTTATC	ATCACTTTAA	AAATAAAAAAA	3450
	CAATTACTCA	GTGCTGTAA	TAAGCAGCAA	TTAATTATGA	TTGATGCCTA	3500
50	CATCACAACA	AAAATGATT	TAACAAATGG	TTGGTCTGCC	TTAGAAAGTA	3550
	TATTTGAACA	TTATCTTGAT	TATATTATTG	ATAATAATAA	AAACCTTATC	3600
	CCTATCCAAG	AAAGTGTGCC	TATCATGGT	TGGAATGAAC	TTGAAAAAAA	3650
	TTAGCCTTGA	ATACATTACT	GGTAAGGTTA	ACGCCATTGT	CAGCAAATTG	3700
	ATCCAAGAGA	ACCAACTTAA	AGCTTCTGTG	ACCGAATGTT	AATTCTCGTT	3750
55	GACCCCTGAGC	ACTGATGAAT	CCCTCTATGA	TTTGGTAA	AATCATTAAG	3800
	TTAAGGTGGA	TACACATCTT	GTCATATGAT	CCCGGTAATG	TGAGTTAGCT	3850
	CACTCATTAG	GCACCCCCAGG	CTTACACTT	TATGCTTCCG	GCTCGTATGT	3900
	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTC	ACACAGGAAA	CAGCTATGAC	3950
	CATGATTACG	CCAAGCGCGC	AATTAACCC	CACTAAAGGG	AACAAAAGCT	4000
60	GGAGCTCCAC	CGCGGTGGCG	GCCGCTCTAG	AACTAGTGG	TCCCCCGGGC	4050
	ATCAGATTTGG	CTATTGGCCA	TTGCATACGT	TGTATCCATA	TCATAATATG	4100

	TACATTTATA	TTGGCTCATG	TCCAACATTA	CCGCCATGTT	GACATTGATT	4150
	ATTGACTAGT	TATTAATAGT	AATCAATTAC	GGGGTCATTA	GTTCATAGCC	4200
	CATATATGGA	GTTCGCGTT	ACATAACCTA	CGGTAAATGG	CCCGCCTGGC	4250
5	TGACCGCCCA	ACGACCCCCG	CCCATTCGACG	TCAATAATGA	CGTATGTTCC	4300
	CATAGTAACG	CCAATAGGGA	CTTCCATTG	ACGTCATGG	GTGGAGTATT	4350
	TACGGTAAAC	TGCCCACCTG	GCAGTACATC	AAGTGTATCA	TATGCCAAGT	4400
	ACGCCCTTA	TTGACGTCAA	TAACGGTAAA	TGGCCCGCT	GGCATTATGC	4450
	CCAGTACATG	ACCTTATGGG	ACTTCTCTAC	TTGGCAGTAC	ATCTACGTAT	4500
10	TAGTCATCGC	TATTACCATG	GTGATGCGGT	TTGGCAGTA	CATCAATGGG	4550
	CGTGGATAGC	GGTTTGACTC	ACGGGGATT	CCAAGTCTCC	ACCCCATTGA	4600
	CGTCAATGGG	AGTTTGTTT	GGCACCAAAA	TCAACGGGAC	TTTCCAAAAT	4650
	GTCGTAACAA	CTCCGCCCCA	TTGACGCAA	TGGGCGGTAG	GCCTGTACGG	4700
	TGGGAGGTCT	ATATAAGCAG	AGCTCGTTA	GTGAACCGTC	AGATCGCCTG	4750
	GAGACGCCAT	CCACGCTGTT	TTGACCTCCA	TAGAAGACAC	CGGGACCGAT	4800
15	CCAGCCTCCG	CGGCCGGGAA	CGGTGATTG	GAACCGGAT	TCCCCGTGCC	4850
	AAAGAGTACG	TAAGTACCGC	CTATAGACTC	TATAGGCACA	CCCCTTGGC	4900
	TCTTATGAT	GCTTACTGT	TTTTGCTTG	GGGCCTATAC	ACCCCCGCTT	4950
	CCTTATGCTA	TAGGTGATGG	TATAGCTTAG	CCTTATAGGTG	TGGGTTATTG	5000
20	ACCATTATTG	ACCACTCCCC	TATTGGTAC	GATACTTTCC	ATTACTAATC	5050
	CATAACATGG	CTCTTGCAC	CAACTATCTC	TATTGGCTAT	ATGCCAATAC	5100
	TCTGTCCTTC	AGAGACTGAC	ACGGACTCTG	TATTTTACA	GGATGGGGTC	5150
	CCATTATTA	TTTACAAATT	CACATATACA	ACAACGCCGT	CCCCCGTGC	5200
	CGCAGTTTTT	ATTAACACATA	GCCTGGGATC	TCCACCGCAA	TCTCGGGTAC	5250
	GTGTTCCGGA	CATGGGCTCT	TCTCCGGTAG	CGGGCGGAGCT	TCCACATCCG	5300
25	AGCCCTGGTC	CCATGCCCTCC	ACGGCCTCAT	GGTCGCTCG	CAGCTCCTTG	5350
	CTCCCTAACAG	TGGAGGGCAG	ACTTAGGCAC	AGCACAAATGC	CCACCAAC	5400
	CAGTGTGCCG	CACAAGGCCG	TGGCGTAGG	GTATGTGTCT	GAAAATGAGC	5450
	GTGGAGATTG	GGCTCGCACG	GCTGACGAG	ATGGAAGACT	TAAGGCAGCG	5500
	GCAGAAGAAG	ATGCAGGCGAG	CTGAGTTGTT	GTATTCTGTAT	AAGAGTCAGA	5550
30	GGTAACTCCC	GTGCGGTGC	TGTTAACGGT	GGAGGGCAGT	GTAGTCTGAG	5600
	CAGTACTCGT	TGCTGCCCG	CGGCCACCA	GACATAATAG	CTGACAGACT	5650
	AACAGACTGT	TCCTTCCAT	GGGTCTTTTC	TGCACTCACC	GTCTCGCGAC	5700
	AGGGATCCAC	CGGTCGCCAC	CATGGTGC	TCCCTCCAAGA	ACGTCACTCAA	5750
	GGAGTTCATG	CGCTTCAAGG	TGCCCATGGA	GGGCACCGTG	ACGGCCACG	5800
35	AGTTCGAGAT	CGAGGGCAG	GGCGAGGGCC	GCCCCCTACGA	GGGCCACAAAC	5850
	ACCGTGAAGC	TGAAGGTGAC	CAAGGGCGGC	CCCCCTGCCCT	TGCCTGGGA	5900
	CATCCTGTCC	CCCCAGTTCC	AGTACGGCTC	CAAGGTGTAC	GTGAAGCACC	5950
	CCGCCGACAT	CCCCGACTAC	AAGAACGCTGT	CCTTCCCCGA	GGGCTTCAAG	6000
	TGGGAGCGCG	TGATGAACCT	CGAGGACGGC	GGCGTGGTGA	CCGTGACCCA	6050
40	GGACTCCTCC	CTGCAGGACG	GCTGCTTCAT	CTACAAGGTG	AAGTCATCG	6100
	GGCGTGAACCT	CCCCCTCCGAC	GGCCCCGTAA	TGCAAGAAGAA	GACCATGGGC	6150
	TGGGAGGCCT	CCACCGAGCG	CCTGTACCCC	CGCGACGGCG	TGCTGAAGGG	6200
	CGAGATCCAC	AAGGCCCTGA	AGCTGAAGGA	CGGGCGGAC	TACCTGGTGG	6250
45	AGTTCAAGTC	CATCTACATG	GCCAAGAACG	CCGTGCACT	CCCCGGCTAC	6300
	TACTACGTGG	ACTCCTAACG	GGACATCACC	TCCCACAACG	AGGACTACAC	6350
	CATCGTGGAG	CACTACGAGC	GCACCCAGGG	CCGGCACCC	CTGTTCTGT	6400
	AGCGGCCGCG	ACTCTAGATC	ATAATCAGCC	ATACCACATT	TGTAGAGGTT	6450
	TTACTTGCTT	AAAAAAACCT	CCACACCTC	CCCCCTGAACC	TGAAACATAA	6500
50	AATGAATGCA	ATTGTTGTG	TAAACTGTT	TATTCAGCT	TATAATGGTT	6550
	ACAAATAAAG	CAATAGCATC	ACAAATTCA	CAAATAAAGC	ATTTTTTCA	6600
	CTGCATTCTA	GTGTTGGCCC	GGGCTGCGAG	AATTGATAT	CAAGCTTATC	6650
	GATACCGCTG	ACCTCGAGGG	GGGGCCCGGT	ACCCAATTG	CCCTATAGTG	6700
	AGTCGTATTA	CGCGCGCTCA	CTGGCCGTG	TTTACAACG	TGCTGACTGG	6750
55	GAAAACCTG	CGGTTACCA	ACTTAATCGC	CTTGCAGCAC	ATCCCCCTTT	6800
	CGCCAGCTGG	CGTAATAGCG	AAAGGCCCG	CACCGATCGC	CCTTCCCAAC	6850
	AGTTGCGCAG	CCTGAATGGC	GAATGAAAT	TGTAAGCGTT	AATATTTGT	6900
	TAAAATTGCG	GTTAAATTT	TGTTAAATCA	GCTCATTTT	TAACCAATAG	6950
	GCCGAAATCG	GCAAAATCCC	TTATAAAATCA	AAAGAATAGA	CCGAGATAGG	7000
60	GTTGAGTGT	GTTCCAGTT	GGAACAAAGAG	TCCACTATTA	AAGAACGTGG	7050
	ACTCCAACGT	CAAAGGGCGA	AAAACCGTCT	ATCAGGGCGA	TGGCCCACTA	7100
	CTCCGGGATC	ATATGACAAG	ATGTGTATCC	ACCTTAACCTT	AATGATTTT	7150

	ACCAAAATCA TTAGGGGATT CATCAGTGCT CAGGGTCAAC GAGAATTAAC 7200
	ATCCGTCAG GAAAGTTAT GATGATGATG TGCTTAAAAA CTTACTCAAT 7250
	GGCTGGTTAT GCATATGCCA ATACATGCGA AAAACCTAA AGAGCTTGCC 7300
5	GATAAAAAG GCCAATTAT TGCTATTAC CGCGGTTTT TATTGAGCTT 7350
	GAAAGATAAA TAAAATAGAT AGGTTTATT TGAAGCTAA TCTTCTTTAT 7400
	CGTAAAAAAT GCCCCTTGG GTTATCAAGA GGGTCATTAT ATTCGCGGA 7450
	ATAACATCAT TTGGTGACCA AATAACTAAC CACTTGTC CTTGTTACTC 7500
	CCCTGAGCTT GAGGGGTTAA CATGAAGGTC ATCGATAGCA GGATAATAAT 7550
10	ACAGTAAAAC GCTAAACCAA TAATCCAAT CCAGGCCATCC CAAATTGGTA 7600
	GTGAATGATT ATAAAATAAC GCAACAGTA ATGGGCAAT AACACCGGTT 7650
	GCATTGGTAA GGCTCACCAA TAATCCCTGT AAAGCACCTT GCTGATGACT 7700
	CTTTGTTGG ATAGACATCA CTCCCTGTAA TGCAAGTAAA GCGATCCCAC 7750
	CACCAGCCAA TAAAATTAA ACAGGGAAAA CTAACCAACC TTCAGATATA 7800
15	AAACGCTAAAAG AGGCAAATGC ACTACTATCT GCAATAAACG CGAGCAGTAC 7850
	TGCCGTTTTT TCGCCCATTT AGTGGCTATT CTTCCGTGCA CAAAGGCTTG 7900
	GAATACTGAG TGTAAAAGAC CAAGACCCGT AATGAAAAGC CAACCATCAT 7950
	GCTATTCTAC ATCACGATT CTGTAATAGC ACCACACCGT GCTGGATTGG 8000
	CTATCAATGC GCTGAAATAA TAATCAACAA ATGGCATCGT TAAATAAGTG 8050
20	ATGTATAACCG ATCAGCTTTT GTTCCCTTAA GTGAGGGTTA ATTGCGCGCT 8100
	TGGCGTAATC ATGGTCATAG CTGTTCCCTG TGTGAAATTG TTATCCGCTC 8150
	ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTAA AAGCCTGGGG 8200
	TGCCTAATGA GTGAGCTAAC TCACATTAAT TGCCTGCGC TCACTGCCCC 8250
	CTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG AATGGCCAA 8300
25	CGCGCGGGGA GAGGCGGTTT GCGTATTGGG CGCTCTTCGG CTTCCCTCGCT 8350
	CACTGACTCG CTGCGCTCGG TCGTTCCGGT GCGGCAGCG CTATCAGCTC 8400
	ACTCAAAGGC GTTAATACGG TTATCCACAG AATCAGGGGA TAACCGAGGA 8450
	AAGAACATGT GAGCAAAGG CCAGCAAAGG GCGAGGAACC GTAAAAGGGC 8500
	CGCGTTGCTG GCGTTTTCC ATAGGCTCGG CCCCCCTGAC GAGCATCACA 8550
30	AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAAGA 8600
	TACCAGCGT TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC 8650
	CCTGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG 8700
	CGCTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT 8750
	CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAAGC CCGACCGCTG 8800
35	CGCCCTATCC GTTAACATC GTCTTGAGTC CAAACCGGTA AGACACGACT 8850
	TATGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT 8900
	GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCTTAACT ACGGCTACAC 8950
	TAGAAGGACA GTATTTGGTA TCTGCCTCT GCTGAAGCCA GTTACCTTCG 9000
	GAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACAC CGCTGGTAGC 9050
40	GGTGGTTTTT TTGTTGCAA GCAGCAGATT ACGCGCAGAA AAAAGGATC 9100
	TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG 9150
	AAAACTCACG TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC 9200
	ACCTAGATCC TTTTAAATTA AAAATGAAGT TTTAAATCAA TCTAAAGTAT 9250
	ATATGAGTAA ACTTGGTCTG ACAGTACCA ATGCTTAATC AGTGAGGGCAC 9300
45	CTATCTACG GATCTGTCTA TTTCGTTCAT CCATAGTGC CTGACTCCCC 9350
	GTCGTGTAGA TAACTACGGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC 9400
	TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA 9450
	TAAACCAGCC AGCCGGAAGG GCGCAGCGA GAAGTGGTCC TGCAACTTTA 9500
	TCCGCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG 9550
50	TTCGCCAGTT AATAGTTTC GCAACGTTGT TGCCATTGCT ACAGGCATCG 9600
	TGGTGTACG CTCGTCGTTT GGTATGGCTT CATTCACTC CGGTTCCCA 9650
	CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG 9700
	CTCTCTCGGT CCTCCGATCG TTGTGAGAAG TAAGTTGGCC CGAGTGGTTAT 9750
	CACTCATGGT TATGGCAGCA CTGCATAATT CTCTTACTGT CATGCCATCC 9800
55	GTAAGATGCT TTTCTGTGAC TTGTGAGTAC TCAACCAAGT CATTCTGAGA 9850
	ATAGTGTATG CGCGGACCGA GTTGCCTTG CCGCGCTCA ATACGGATA 9900
	ATACCGCGCC ACATAGCAGA ACTTTAAAG TGCTCATCAT TGGAAAACGT 9950
	TCTTCGGGGC GAAAACCTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC 10000
	GATGTAACCC ACTCGTGAC CCAACTGATC TTCAAGCATCT TTTACTTTCA 10050
60	CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAATGC CGCAAAAAAG 10100
	GGAATAAGGG CGACACGAA ATGTTGAATA CTCATACTCT TCCCTTTTCA 10150
	ATATTATTGAGA AGCATTATC AGGGTTATTG TCTCATGAGC GGATACATAT 10200

TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC 10250
 CGAAAAGTGC CAC 10263

5 SEQ ID NO:3 (PTnMod (Oval/Red) Chicken)

CTGACGCCCT CTGTAGCCGC GCATTAAGCG CGCCGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCCTTCGC 100
 TTTCTTCCTC TGCTTCTCG CCACGTTCGC CGGCATCAGA TTGGCTATTG 150
 10 GCCATTGCACT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
 CATGTCCAAC ATTACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA 250
 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG 300
 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
 CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
 15 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAC TGCCCA 450
 CTTGGCAGTA CATCAAGTGT ATCATATGCC AACTACGCC CCTATTGACG 500
 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCATGA CATGACCTTA 550
 TGGGACTTTC CTACTGGCA GTACATCTAC GTATTAGTC TCGCTATTAC 600
 CATGGTGTG CCGTTTGGC AGTACATCAA TGGCGTGGA TAGCGGTTG 650
 20 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTG 700
 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC 750
 CCCATTGACG CAAATGGCGC GTAGGGCTGT ACGGTGGGAG GTCTATATAA 800
 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
 TGTTTGACCC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG 900
 25 GGAACGGTGC ATTGGAACGC GGATTCCCCG TCCAAGAGT GACGTAAGTA 950
 CGCCTATAG ACTCTATAGG CACACCCCT TGCTCTTAT GCATGCTATA 1000
 CTGTTTTGG CTTGGGGCCT ATACACCCCT GCTCCCTTAT GCTATAGGTG 1050
 ATGGTATAGC TTAGCCTATA GGTGTGGTT ATTGACCTT ATTGACCACT 1100
 CCCCTATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTT 1150
 30 GCCACAACTA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTCAGAGAC 1200
 TGACACGGAC TCTGTATTT TACAGGATGG GGTCCCATTT ATTATTTACA 1250
 AATTCAACATA TACAACAACG CCGTCCCCCG TGCCCGCAGT TTTTATTAAA 1300
 CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTG CGGACATGGG 1350
 CTCTCTCCGG GTAGCGGGCG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
 35 CTCCAGCGGC TCATGGTCGC TCGGCAGCTC CTGCTCTTA ACAGTGGAGG 1450
 CCAGACTTAG GCACAGCACA ATGCCCCACA CCACAGCTGT GCGCACAAG 1500
 GCGGTGGCGG TAGGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGCTCG 1550
 CACGGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG 1600
 GCAGCTGAGT TGTTGTATTC TGATAAGAGT CAGAGGTAAC TCCCCTTGCG 1650
 40 GTGCTGTTAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TCGTTGCTGC 1700
 CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTTT 1750
 CCATGGGTCT TTTCTGCACT CACCGTCGGA CCATGTGTGA ACTTGATATT 1800
 TTACATGATT CTCTTACCA ATTCTGCCCC GAATTACACT TAAAACGACT 1850
 45 CAACAGCTTA ACGTTGGCTT GCCACGCATT ACTGACTGT AAAACTCTCA 1900
 CTCTTACCGA ACTTGGCCGT AACCTGCCA CCAAAGCGAG AACAAAACAT 1950
 AACATCAAC AACATCGACCG ATTGTTAGGT AATCGTCACC TCCACAAAGA 2000
 GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC 2050
 GATGCCATT GTACTTGGTG ACTGGTCTGA TATTCTGTGAG CAAAACGAC 2100
 TTATGGTATT GCGAGCTTCA GTCGCACTAC ACGGTGTTG TGTTACTCTT 2150
 50 TATGAGAAAG CGTTCCCGCT TTCAAGCAA TGTTCAAAGA AAGCTCATGA 2200
 CCAATTCTA GCGACCTTG CGACCATCT ACCGAGTAAC ACCACACCGC 2250
 TCATTGTCAG TGATGTCGC TTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
 AAGCTGGGTG GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
 AGACCTAGGA CGGGAAAAT CCAACCTAT CAGCAACTTA CATGATATGT 2400
 55 CATCTAGTCA CTCAAAGACT TTAGGCCTATA AGAGGCTGAC TAAAAGCAAT 2450
 CCAATCTCAT GCCAAATTCT ATTGTATAAA TCTCGCTCTA AAGGCCGAAA 2500
 AAATCAGCGC TCGACACCGA CTCATTGTCA CCACCCGTCA CCTAAAATCT 2550
 ACTCAGCGTC GGCAAAGGAG CCATGGGTT TAGCAACTAA CTTACCTGTT 2600
 GAAATTGAA CACCCAAACA ACTTTGTTAAT ATCTTATTCGA AGCGAATGCA 2650
 60 GATTGAAGAA ACCTTCCGAG ACTTGAAAAG TCCTGCTTAC GGACTAGGCC 2700
 TACGCCATAG CCGAACGAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750

ATCGCCCTGA TGCTTCACT AACATGTTGG CTTGCGGGCG TTCATGCTCA 2800
 GAAACAAGGT TGGGACAAGC ACTTCAGGC TAAACAGTC AGAAATCGAA 2850
 ACGTACTCTC AACAGTTCGC TTAGGCATGG AAGTTTGCG GCATTCTGGC 2900
 TACACAATAA CAAGGGAAAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 5 AAATTTATTC ACACATGGTT ACGTTGGG GAAATTATGA TAATGATCCA 3000
 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 TTTTTGTGG ATCTGCTGT CCTTCTAGTT GCCAGCCATC TGTTGTTTGC 3100
 CCCTCCCCCG TGCCCTCTT GACCTGGAA GGTGCCACTC CCACTGTCCT 3150
 10 TTCCATAATAA AATGAGGAAA TTGCATCGCA TTGCTCTGAGT AGGTGTCATT 3200
 CTATTCTGGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTGGGAA 3250
 GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCGGTAC CTCTCTCTCT 3350
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAAGAA 3400
 TTGACCCGGT GACCAAAGGT GCCTTTATC ATCACTTTAA AAATAAAAAA 3450
 15 CAATTACTCA GTGCCTGTTA TAAGCAGCAA TTAATTATGA TTGATGCCTA 3500
 CATCACAAACA AAAACTGATT TAACAAATGG TTGGCTGCGC TTAGAAAGTA 3550
 TATTGAAACA TTATCTTGAT TATATTATTAA ATAATAATAA AAACCTTATC 3600
 CCTATCAGAAGT GATGATGCC TATCATTGGT TGAATGAAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GTAAAGGTAA ACGCATTGT CAGCAAATTG 3700
 20 ATCCAAGAGA ACCAACTTAA AGCTTCTGT ACGGAATGTT AATTCTCGTT 3750
 GACCCCTGAGC ACTGATGAAT CCCCTAATGA TTTGGTAAA AATCATTAAG 3800
 TTAAGGTGGA TACACATTT GTCATATGAT CCCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAA CAGCTATGAC 3950
 25 CATGATTACG CCAAGCGCG AATTAACCCCT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCGCCTCTAG AACTAGTGGA TCCCCCGGGG 4050
 AGGTCAAGAT GGTTTCTTCA CTGTTGTCA ATTCTTATTAT TTCAATACAG 4100
 ACAATAGCT TCTATAACTG AAATATTTT GCTATTGTAT ATTATGATTG 4150
 TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTTCACAA TTCCCTCTGTC 4200
 30 ATCTGCCAGG CCATTAAGTT ATTCATGGAA GATCTTGTGAG GAACACTGCA 4250
 AGTTCATATC ATAAACACAT TTGAAATTGA GTATTGTTT GCATTGTATG 4300
 GAGCTATGTT TTGCTGTATC CTCAGAAAAA AAGTTGTTA TAAAGCATTC 4350
 ACACCCATAA AAAGATAGAT TAAATATTTC CAGCTATAGG AAAGAAAGTG 4400
 CGTCTGCTCT TCACTCTAGT CTCAGTTGGC TCCTTCACAT GCATGCTCT 4450
 35 TTATTTCTCC TATTTGTCA AGAAAATAAT AGGTCACTGTC TTGTTCTCAC 4500
 TTATGCTCTG CCTAGCATGG CTCAGATGCA CGTCTGTAGAT ACAAGAAGGA 4550
 TCAAATGAAA CAGACTCTGTG GTCTGTACT ACAACCATAG TAATAAGCAC 4600
 ACTAACTAAT AATTGCTAAT TATGTTTCC ATCTCTAAGG TTCCCACATT 4650
 TTTCTGTTT CTTAAAGATC CCATTATCTG GTTGTAACTG AAGCTCAATG 4700
 40 GAACATGAGC AATATTTCCC AGTCTTCTCT CCCATCCAAC AGTCCTGTATG 4750
 GATTAGCAGA ACAGGCAGAA AACACATTGT TACCCAGAAT TAAAAACTAA 4800
 TATTGCTCT CCATTCAATC CAAAATGGAC CTATTGAAAC TAAAATCTAA 4850
 CCCAATCCCA TAAATGATT TCTATGGCGT CAAAGTCAA ACTTCTGAAG 4900
 GGAACCTGTG GGTGGTCAC AATTCAAGGCT ATATATTCCC CAGGGCTCAG 4950
 45 CGGATCTCA TGGGCTCCAT CGGTGCAGCA AGCATGGAAT TTTGTTTGA 5000
 TGTATTCAAG GAGCTCAAAG TCCACCATGC CAATGAGAAC ATCTTCTACT 5050
 GCCCCATTGC CATCATGTCA GCTCTAGCCA TGTTATACCT GGGTGCAGAA 5100
 GACAGCACCA GGGAAATTCTGT GCGCTCTCC AAGAACGTCA TCAAGGAGTT 5150
 CATCGCCTTC AAGGTGCGCA TGGAGGGCAC CGTGAACGGC CACGAGTTCG 5200
 50 AGATCGAGGG CGAGGGCGAG GGCGCCCCCTT ACAGGGGCCA CAACACCGTG 5250
 AAGCTGAAGG TGACCAAGGG CGGCCCCCTG CCCTTCGCGCT GGGACATCCT 5300
 GTCCCCCCAG TTCCAGTAGC GCTCCAAAGGT GTACGTGAAG CACCCCGCCG 5350
 ACATCCCCGA CTACAAGAAG CTGTCCTTCC CCGAGGGCTT CAAGTGGGAG 5400
 CGCGTGTGTA ACTTCGAGGA CGCGGGCGTG GTGACCGTGA CCCAGGACTC 5450
 55 CTCCCTGCAG GACGGCTGCT TCATCTACAA GGTGAAGTTC ATCGGCGTGA 5500
 ACTTCCCCCTC CGACGGCCCC GAAATGAGA AGAAGACCAT GGGCTGGGAG 5550
 GCCTCCACCG AGCGCCTGTA CCCCCCGCAG GGCCTGCTGA AGGGCGAGAT 5600
 CCACAAGGCC CTGAAGCTGA AGGACGGCGG CCAACTACCTG GTGGAGTTCA 5650
 AGTCCATCTA CATGGCCAAG AAGCCCGTGC AGCTGCCCGG CTACTACTAC 5700
 60 GTGGACTCCA AGCTGGACAT CACCTCCCCAC AACGAGGACT ACACCATCGT 5750
 GGAGCACTAC GAGCCGACCG AGGGCCGCCA CCACCTGTTC CTGTAGCGGC 5800

CGCGACTCTA GATCATAATC AGCCATACCA CATTGTAGA GGTTTACTT 5850
 GCTTTAAAAA ACCTCCCACA CCTCCCCCTG AACCTGAAAC ATAAAATGAA 5900
 TGCAATTGTT GTTGTAACT TGTTTATTGC AGCTTATAAT GGTTACAAAT 5950
 AAAGCAATAG CATCACAAAT TTCACAAATA AAGCATTTTT TTCACTGCAT 6000
 5 TCTAGTTGTG GTCGAGAAG GGCAGATTCT GCAGATATCC ATCACACTGG 6050
 CGGCCGCTCG AGGGGGGCC CGGTACCCAA TTGCGCCCTAT AGTGAGTCGT 6100
 ATTACCGCGC CTCACTGCCG GTCGTTTAC AACGTCGTGA CTGGGAAAC 6150
 CCTGGCGTTA CCCCACCTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG 6200
 CTGGCGTAAT AGCGAAGAGG CCCGACCCGA TCGCCCTTCC CAACAGTTGC 6250
 10 GCAGCCTGAA TGGCAATGG AAATTGTAAG CGTTAATATT TTGTTAAAAT 6300
 TCGCGTTAAA TTTTGTTAA ATCAAGCTCAT TTTTTAACCA ATAGGCCGAA 6350
 ATCGGCAAA TCCCTTATAA ATCAAAAGAA TAGACCGAGA TAGGGTTGAG 6400
 TGTTGTTCCA GTTGGAAACA AGAGTCCACT ATTAAAGAAC GTGGACTCCA 6450
 ACGTCAAAGG GCGAAAACC GTCTATCAGG GCGATGGCCC ACTACTCCGG 6500
 15 GATCATATGA CAAGATGTGT ATCCACCTTA ACTTAATGAT TTTTACCAAA 6550
 ATCATTAGGG GATTCATCAG TGCTCAGGGT CAACGAGAAT TAACATTCCG 6600
 TCAGGGAAAGC TTATGATGAT GATGTGCTTA AAAACTTACT CAATGGCTGG 6650
 TTATGCATAT CGCAATACAT CGGAAAACC TAAAAGAGCT TGCCGATAAA 6700
 AAAGGCCAAT TTATTGCTAT TTACCGCGC TTTTTATTGA GCTTGAAAGA 6750
 20 TAAATAAAAT AGATAGGTT TATTGAAAGC TAAATCTCT TTATCGTAAA 6800
 AAATGCCCTC TTGGGTTATC AAGAGGGTCA TTATATTTCG CGGAATAACA 6850
 TCATTTGGTG ACGAAATAAC TAAGCACTTG TCTCCTGTTT ACTCCCCTGA 6900
 GCTTGAGGGG TTAACATGAA GGTCAATCGAT AGCAGGATAA TAATACAGTA 6950
 AAACGCTAAA CCAATAATCC AAATCCAGCC ATCCCAAATT GTAGTGAAT 7000
 25 GATTATAAAAT AACAGCAAC ACTAATGGGC CAATAACACC GGTTGCATTG 7050
 GTAAGGCTCA CCAATAATCC CTGTAAGCA CCTGCTGTAT GACTCTTGT 7100
 TTGGATAGAC ATCACTCCCT GTAATGCAGG TAAAGCGATC CCACCACAG 7150
 CCAATAAAAT TAAAACAGGG AAAACTAACC AACCTTCAGA TATAAACGCT 7200
 AAAAAGGCAA ATGCACTACT ATCTGCAATA AATCCAGCA GTACTGCCGT 7250
 30 TTTTCGCC CATTAGTGG CTATTCTTC TGCCACAAAG GCTTGAATA 7300
 CTGAGTGTAA AAGACCAAGA CCCGCTAATG AAAAGCCAAC CATCATGCTA 7350
 TTCCATCCAA AACGATTTCG GGTAAATAGC ACCCACACCG TTGCGGGAAAT 7400
 TTGGCCTATC AATTGCGCTG AAAAATAAT AATCAACAAA ATGGCATCGT 7450
 TTTAAATAAA GTGATGTATA CCGAATTTCAG CTTTGTCTC CTTTAGTGAN 7500
 35 GTGAAATTGCG CCGCTTGGCG TAATCATGGT CATAGCTGTT TCCTGTTG 7550
 ATTGTTATC CGCTCACAAAT TCCACACAAAC ATACGAGCCG GAAGCATAAA 7600
 GTGTAAAGCC TGGGGTGCCT AATGAGTGAN CTAACTCACA TTAATTGCGT 7650
 TGCGCTCACT GCCCCGTTT CAGTCGGAA ACCTGTCGTG CCAGCTGCAT 7700
 TAATGAATCG GCCAACCGCG GGGGAGAGGC GGTTCGCTA TTGGGCGCTC 7750
 40 TTCCGCTTCC TCGCTCACTG ACTCGCTCG CTCGGCTGTT CGGCTGCCG 7800
 GAGCGGTATC AGCTCACTCA AAGGGTAA TACGGTTATC CACAGAATCA 7850
 GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG 7900
 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCATAGG CTCCGCC 7950
 CTGACGAGCA TCACAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG 8000
 45 ACAGGACTAT AAAGATACCA GGCCTTTCCC CCTGGAAGCT CCCTCGTGC 8050
 CTCTCCGTGTT CCGACCCCTGC CGCTTACCGG ATACCTGTCC GCCTTCTCC 8100
 CTTCGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT 8150
 TCGGGTAGG TCGTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT 8200
 TCAGCCGAC CGCTGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC 8250
 50 CGGTAAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT 8300
 AGCAGAGCGA GTTATGTAGG CGGTGCTACA GAGTTCTGA AGTGGTGGCC 8350
 TAACTACGGC TACACTAGAA GGACAGTATT GGATGTGATC GCTCTGCTGA 8400
 AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA 8450
 55 ACCACCGCTG GTAGCGGTGG TTTTTTGTG TGCAAGCAGC AGATTACGCG 8500
 CAGAAAAAAA GGATCTCAAG AAGATCCTT GATCTTTCT ACGGGGCTG 8550
 ACGCTCAGTG GAACGAAAC TCACGTTAAG GGATTTGGT CATGAGATTA 8600
 TCAAAAAGGA TCTTCACCTA GATCCTTTA AATTAAAAAT GAAGTTTAA 8650
 ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT 8700
 TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTTATTGTT TCATCCATA 8750
 60 GTTGCCTGAC TCCCCGCTCGT GTAGATAACT ACGATACGGG AGGGCTTAC 8800
 ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC 8850

CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT 8900
 GGTCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAAATT GTTGCCTGGGA 8950
 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTGCTCA 9000
 5 TTGCTACAGG CATCGTGGTG TCACCGCTCGT CGTTGGTAT GGCTTCATTC 9050
 AGCTCCGGTT CCCAACGATC AAGGGAGTT ACATGATCCC CCATGTTGTG 9100
 CAAAAAAAGCG GTTAGCTCCT TCGGTCTCC GATCGTTGTG AGAAGTAAGT 9150
 TGGCCGCAGT GTTATCACTC ATGCTTATGG CAGCAGTGC TAATTCTCTT 9200
 ACTGTCATGC CATCGTAAAG ATGCTTTCT GTGACTGGTG AGTACTCAAC 9250
 10 CAAGTCATTC TGAGAATAGT GTATGCGGC ACCGAGTTGC TCTTGCCCCG 9300
 CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGTCTC 9350
 ATCATTGGAA AACGTTCTTC GGGGCGAAA CTCTCAAGGA TCTTACCGCT 9400
 GTTGAGATCC AGTCGATGT AACCCACTCG TGCCACCCAAAC TGATCTTCAG 9450
 CATCTTTAC TTTCACCAGC GTTCTGGGT GAGCAAAAAC AGGAAGGCAA 9500
 15 AATGCCGCAA AAAAGGAAAT AAGGGGACAA CGGAAATGTT GAATACTCAT 9550
 ACTCTCCCTT TTTCATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA 9600
 TGAGCGGATA CATATTGAA TGTATTAGA AAAATAAAACA AATAGGGTT 9650
 CCGCGCACAT TTCCCCGAAA AGTGCAC 9678

20 SEQ ID NO:4 (PTnMod (Oval/Red) Quail)

CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGGGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCCTTCGC 100
 25 TTTCTTCCCT TCCCTTCTCG CCACGTTCGC CGGCATCAGA TTGGCTATTG 150
 GCCATTGCACT ACCTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
 CATGTCCAAAC ATTACCGCA TGTTGACATT GATTATTGAC TAGTTATTAA 250
 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG 300
 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGCGTGACCG CCCAACGACC 350
 CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
 30 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCA 450
 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG 500
 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA 550
 TGGGACTTTTC CTACTTGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
 CATGGTGTAG CGGTTTTGGC AGTACATCAA TGCGCTGGAA TAGCGGTTTG 650
 35 ACTCACGGGG ATTTCGAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTG 700
 TTTGGCACC AAAATCAACG GGACTTCCA AAATGTCGTA ACAACTCCGC 750
 CCCATTGACG CAAATGGCG GTAGGGTGT ACGGTGGGAG GTCTATATAA 800
 GCAGAGCTCG TTTACTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
 TGTTTGACCC TCCATAGAAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG 900
 40 GGAACGGTGC ATTGGAAACGC GGATCCCCG TGCCAAGAGT GACGTAAGTA 950
 CCCCTATAG ACTCTATAGG CACACCCCTT TGCGCTCTTAT GCATGCTATA 1000
 CTGTTTTGG CTTGGGGCCT ATACACCCCC GCTTCCCTTAT GCTATAGGTG 1050
 ATGGTATAGC TTAGCCTATA GGTGTGGTT ATTGACCAATT ATTGACCACT 1100
 45 CCCCTATTGG TGACGTAATTCTCCATTAAAC ATGGCTCTTT 1150
 GCCACAACTA TCTCTATTGG CTATATGCC ATACTCTGTC CTTCAGAGAC 1200
 TGACACGGAC TCTGTATTTC TACAGGATGG GGTCCCATT ATTATTTACA 1250
 AATTACACATA TACAACAACG CCGTCCCCCG TGCGCGCAGT TTTTATTAAA 1300
 CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTG CGGACATGGG 1350
 50 CTCTTCTCCG GTAGCGGGGG AGCTTCCACA TCCGAGCCCT GGTCCCAGTC 1400
 CTCCAGCGGC TCATGCTCGC TCGGCAGCTC CTGCTCTCTA ACAGTGGAGG 1450
 CCAGACTTAG GCACAGCACA ATGCCACCA CCACCACTGT GCGGCACAAG 1500
 GCGCGTGGCG TAGGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGGCTCG 1550
 CACGGCTGAC GCAGATGGAA GACTTAAGGG AGCGGCAGAA GAAGATGCAG 1600
 GCAGCTGAGT TGTGTATTG TGATAAGAGT CAGAGGTAAC TCCCGTTGCG 1650
 55 GTGCTGTTAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TGTTGCTGC 1700
 CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTTT 1750
 CCATGGGTCT TTTCTGCACT CACCGTCGGA CCATGTTGTA ACTTGATATT 1800
 TTACATGATT CTCTTTACCA ATTCTGCCCT GAATTACACT TAAACGACT 1850
 CAACAGCTTA ACGTTGGCCTT GCCACGATT ACTTGACTGT AAAACTCTCA 1900
 60 CTCTTACCGA ACTTGGCCGT AACCTGCCAA CCAAAGCGAG AACAAAACAT 1950
 AACATCAAAC GAATCGACCG ATTGTTAGGT AATCGTCACC TCCACAAAGA 2000

	GGCACTCGCT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TCGGGAATAC	2050
	GATGCCATT	GTACTTGTG	ACTGGTCTGA	TATTCGTGAG	CAAAAACGAC	2100
5	TTATGGTATT	GCGAGCTTCA	CTCGCACTAC	ACGGTCGTT	TGTTACTCTT	2150
	TATGAGAAAG	CGTTCGGCT	TTCAGAGCAA	TGTTCAAAGA	AAGCTCATGA	2200
10	CCAATTCTCA	GCGACCTTG	CGAGCATTCT	ACCGAGTAAC	ACCACACCGC	2250
	TCATTGTCAG	TGATGCTGGC	TTTAAAGTGC	CATGGTATAA	ATCCGTTGAG	2300
	AAGCTGGTT	GGTACTGTT	AACTCGAGTA	AGAGGAAAG	TACAATATGC	2350
	AGACCTAGGA	GCGGAAACT	GGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
15	CATCTAGTCA	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAAGCAAT	2450
20	CCAATCTCAT	GCCAAATTCT	ATTGTATAAA	TCTCGCTCA	AAGGCCGAAA	2500
	AAATCAGCG	TCGACACGGA	CTCATGTC	CCACCCGTC	CCTAAAATCT	2550
	ACTCAGCGTC	GGCAAAGGAG	CCATGGGTT	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTGCAA	CACCCAAACA	ACTTGTTAAT	ATCTATTGCA	AGCGAATGCA	2650
25	GATTGAAGAA	ACCTTCGGAG	ACTTGAAAAG	TCTGCGCTAC	GGACTAGGCC	2700
	TACGCCATAG	CCGAACGAGC	AGCTCAGAGC	GTTCATGATAT	CATGCTGCTA	2750
	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	CTTGCAGGCC	TTCATGCTCA	2800
	GAAACAAGGT	TGGGACAAGC	ACTTCAGGC	TAACACAGTC	AGAAATCGAA	2850
30	ACGTACTCTC	AACAGTCGCG	TTAGGCATGG	AAAGTTTGC	GCATTCTGGC	2900
	TACACAATAA	CAAGGGAAAGA	CTTACTCGTG	GCTGCAACCC	TACTAGCTCA	2950
35	AAATTATTC	ACACATGTT	ACGCTTGGG	GAAATTATGA	TAATGATCCA	3000
	GATCACTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTGTGTTGG	3050
	TTTTTGTGG	ATCTGCTGTG	CCTCTAGTT	GCCAGCCATC	TGTTGTTGTC	3100
	CCCTCCCCCG	TGCCTCCTT	GACCTGGAA	GGTGCACACT	CCACTGTCCT	3150
40	TTCTCTATAA	AATGAGGAA	TTGCATCGCA	TTGCTCTGAGT	AGGTGTCATT	3200
45	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACA	GCAAGGGGA	GGATTGGGAA	3250
	GACAATAGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAAGG	TGCTGAAGAA	3400
50	TTGACCCGGT	GACCAAGGAT	GCTTTTATC	ATCACTTTAA	AAATAAAAAA	3450
55	CAATTACTCA	GTGCCTGTTA	TAAGCAGCAA	TTAATTATGA	TTGATGCCA	3500
	CATCACAAACA	AAAACGTATT	TAACAAATGG	TTGGTCTGCC	TTAGAAAGTA	3550
	TATTTGAACA	TTATCTTGAT	TATATTATTG	ATAATAATAA	AAACCTTATC	3600
	CCTATCCAAG	AAAGTGTGCC	TATCATTGGT	TGGAATGAAC	TTGAAAAAAA	3650
60	TTAGCCTTGA	ATACATTACT	GGTAAGGTAAC	ACGCCATTGT	CAGCAAATTG	3700
	ATCCAAGAGA	ACCAACTAA	AGCTTCCTG	ACGGAATGTT	AATTCTCGTT	3750
	GACCCCTGAGC	ACTGATGAAT	CCCCTAATGA	TTTGGTAA	AATCATTAAAG	3800
	TTAAGGTGGA	TACACATCTT	GTCATATGAT	CCCCGTAATG	TGAGTTAGCT	3850
	CACTCATTAG	GCACCCCAAG	CTTACACTT	TATGCTTCG	GCTCGTATGT	3900
	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTC	ACACAGGAA	CAGCTATGAC	3950
	CATGATTACG	CCAAGCGCGC	ATTAAACCT	CACTAAAGGG	AACAAAAGCT	4000
	GGAGCTCCAC	CGCGGTGGCG	GCCGCTCTAG	AACTAGTGG	TCCCCCGGGG	4050
	AGGTCAAGAT	GGTTTCTTA	CTGTTGTCA	ATTCTATTAT	TTCAATACAG	4100
	AACAAAAGCT	TCTATAACTG	AAATATATT	GCTATTGTAT	ATTATGATTG	4150
	TCCCTCGAAC	CATGAACACT	CCTCAGCTG	AAATTCAAA	TTCCCTCTGTC	4200
65	ATCTGCCAGG	CTGGAAGATC	ATGGAAGATC	TCTGAGAAC	ATTGCAAGTT	4250
	CATACCAATAA	ACTCATTGG	AATTGAGTAT	TATTTTGCTT	TGAATGGAGC	4300
	TATGTTTGCG	AGTTCCCTCA	GAAGAAAAGC	TTGTTATAAA	GCGTCTACAC	4350
	CCATCAAAAG	ATATATTAA	ATATTCAAC	TACAGAAAGA	TTTTGTCG	4400
	TCTTCACTCT	GATCTCAGTT	GGTTTCTCA	CGTACATGCT	TCTTATTG	4450
70	CCTATTCTGT	CAAGAAAATA	ATAGGTCAAG	TCTGTTCTC	ACTTATCTCC	4500
	TGCCTAGCAT	GGCTTAGATG	CACTGTCAC	ATTCAAGAAG	GATCAAATGA	4550
	AACAGACTTC	TGGTCTGTTA	CAACAAACCAT	AGTAATAAAC	AGACTAACTA	4600
	ATAATTGCTA	ATTATGTTT	CCATCTCTA	GGTCTCCAC	TTTTTCTGTT	4650
	TTAAGATCCC	ATTATCTGGT	TGTAAGTGA	GCTCAATGGA	ACATGAACAG	4700
75	TATTTCTCTAG	TCTTTCTCC	AGCAATCTG	ACGGGATTAGA	AGAACTGGCA	4750
	GAAAACACTT	TGTTACCCAG	AATTAAAAC	TAATAATTG	TCTCCCTCA	4800
	ATCCAAATG	GACCTATTGA	AACTAAAATC	TGACCCAATC	CCATTAAAT	4850
	ATTCTATGG	CGTCAAAGGT	CAAACCTTTG	AAAGGAACCT	GTGGGTGGGT	4900
	CCCAATTCTAG	GCTATATATT	CCCCCAGGGCT	CAGCGGATCT	CCATGGGCTC	4950
	CTCGTGCAGC	AAGCATGAA	TTTGCCTTG	ATGTATTCAA	GGAGCTCAA	5000
	GTCCACCATG	CCAATGACAA	CATGCTCTAC	TCCCCCTTG	CCATCTGTCA	5050

ACTCTGGCCA TGGTCTCCCT GGGTGCAGGG GACAGGCCA GGGATTGCGT 5100
 GCGCTCCTCC AAGAACGTCA TCAAGGAGTT CATGCGCTTC AAGGTGCGCA 5150
 TGGAGGGCAC CGTGAACGGC CACGAGTTG AGATCGAGGG CGAGGGCGAG 5200
 5 GGCCGCCCTC ACGAGGGCCA CAACACCGTG AAGCTGAAGG TGACCAAGGG 5250
 CGGCCCCCTG CCCTTCGCTT GGGACATCCT GTCCCCCAG TTCCAGTACG 5300
 GCTCCAAGGT GTACGTGAAG CACCCCGCCG ACATCCCCGA CTACAAGGAAG 5350
 CTGTCCCTCC CCGAGGGCTT CAAGTGGGAG CGCGTGTGAG ACTTCGAGGA 5400
 CGGCAGGGCTG GTGACCGTGA CCCGAGACTC CTCCCTGCGAG GACGGCTGCT 5450
 TCATCTACAA GGTGAAGTT ATCGGCGTGA ACTTCCCTC CGACGGCCCC 5500
 10 GTAATGCAGA AGAACGACAT GGGCTGGGAG GCCTCCACCG AGCGCCTGTA 5550
 CCCCCGCGAC GGCCTGCTGA AGGGCGAGAT CCACAAAGGCC CTGAAGCTGA 5600
 AGGACGGCGG CCACTACCTG GTGGAGTTCA AGTCCATCTA CATGGCCAAG 5650
 AAGCCCGTGC AGCTGCCCGG CTACTACTAC GTGGACTCCA AGCTGGACAT 5700
 CACCTCCAC AACGAGGAAC ACACCATCGT GGAGCAGTAC GAGCGCACCG 5750
 15 AGGGCCGCCA CCACCTGTT CTGTAGCGGC CGCGACTCTA GATCATAATC 5800
 AGCCATACCA CATTGTAGA GGTTTACTT GCTTTAAAAA ACCTCCACA 5850
 CCTCCCTCTG AACCTGAAC ATAAAATGAA TGCAATTGTT GTTGTAACT 5900
 TGTTTATTGC AGCTTATAAT GTTACAACAA AAAGCAATAG CATCACAAT 5950
 TTCACAAATA AAGCATTTTT TTCACTGCAT TCTAGTTGTG GCTCGAGAAG 6000
 20 GGCAGATTCT CGAGATATCC ATCACACTGG CGGCCGCTCG AGGGGGGGCC 6050
 CGGTACCCAA TTGCCCCAT ATGTAGCTGT ATTACCGCGC CTCACTGGCC 6100
 GTCGTTTTAC AACGTCGTGA CTGGAAAAC CCTGGCGTTA CCCAACTTAA 6150
 TCGCCTTGCA GCACATCCCC CTTCGCGAG CTGGCGTAAT AGCGAAGAGG 6200
 CCCGCACCGA TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG 6250
 25 AAATTGTAAG CGTTAATATT TTGTTAAAAT TCGCGTTAAA TTTTTGTTAA 6300
 ATCAGCTCAT TTTTTAACCA ATAGGCCGA ATCGGCAAAA TCCCTTATAA 6350
 ATCAAAAGAA TAGACCGAGA TAGGTTGAG TGGTGTCCA GTTTGGAAACA 6400
 AGAGTCCACT ATAAAGAAC GTGGACTCCA ACCTCAAAGG GCGAAAACC 6450
 GTCTATCAGG GCGATGGCC ACTACTCCGG GATCATATGA CAAGATGTT 6500
 30 ATCCACCTTA ACTTAATGAT TTTTACAAA ATCATTAGGG GATTCACTCAG 6550
 TGCTCAGGGT CAACGAGAAAT TAACATTCCG TCAGGAAAGC TTATGATGAT 6600
 GATGTGCTTA AAAACTTACT CAATGGCTGG TTATGCATAT CGCAATACAT 6650
 GCGAAAACC TAAAAGAGCT TGCGGATAAA AAAGGCCAAT TTATTGCTAT 6700
 TTACCGCGGC TTTTTATGTA GCTTGAAGA TAAATAAAAT AGATAGGTT 6750
 35 TATTGAAAGC TAAATCTCT TTATCGTAAACCA AATGCCCTC TTGGGTTATC 6800
 AAGAGGGTCA TTATATTTCG CGGAATACCA TCATTTGGTG ACGAAATAAC 6850
 TAAGCACTTG TCTCTGTT ACTCCCCCTGA GCTTGAGGGG TTAACATGAA 6900
 GGTATCGAT AGCAGGATAA TAATACAGTA AAACGCTAAA CCAATAATCC 6950
 AAATCCAGCC ATCCCAAATT GGTAGTGAAT GATTATAAAAT AACAGCAAAC 7000
 40 AGTAATGGGC CAATAACACC GGTGCGATTG GTAAGGCTCA CCAATAATCC 7050
 CTGTAAAGCA CCTTGCTGAT GACTTTGT TTGGATAGAC ATCACTCCCT 7100
 GTAATGCAGG TAAAGCGATC CCACCAACCG CCAATAAAAT TAAAACAGGG 7150
 AAAACTAACCC AACCTTCAGA TATAAACGCT AAAAGGCAA ATGCACTACT 7200
 ATCTGCAATA AATCCGAGCA GTACTGCCGT TTTTCGCC CATTAGTGG 7250
 45 CTATTCTTC TGCCACAAAG GCTTGGATAA CTGAGTGTAA AAGACCAAGA 7300
 CCCGCTAATG AAAAGCCAAC CATCATGCTA TTCCATCCAA AACGATTTTC 7350
 GGTAAATAGC ACCCACACCG TTGGGGAAAT TTGGCCTATC AATTGCGCTG 7400
 AAAATAAAAT AATCAACAAA ATGGCATCGT TTAAATAAA GTGATGTATA 7450
 CCGAATTCTAG CTTTTGTTCC CTTTAGTGAG GGTAAATTGC GCGCTTGGCG 7500
 50 TAATCATGGT CATAGCTGTT TCCTGTGTA AATTGTTATC CGCTCACAAT 7550
 TCCACACAAAC ATACGAGGCCG GAAGCATAAA GTGTAAGGCC TGGGGTGCT 7600
 ATGAGTGAG CTAACCTACA TTAATGCGT TGCGCTACT GCGCGCTTTC 7650
 CAGTCGGGAA ACCTGTCGTG CCAGCTGCG TAATGAATCG GCGAACGCGC 7700
 GGGGAGAGGC GGTGGCTA TTGGCGCTC TTCCGCTTCC TCGCTCACTG 7750
 55 ACTCGCTGCCG CTCGGCTGTT CGGCTGCCG GAGCGGTATC AGCTCACTCA 7800
 AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA 7850
 CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCGCGT 7900
 TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAT 7950
 CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA 8000
 60 GGCCTTACCGG CCTGGAAAGCT CCCTCGTGCCT CTCTCCTGTT CCGACCCCTGC 8050
 CGCTTACCGG ATACCTGTCC GCCTTCTCC CTTCGGGAAG CGTGGCGCTT 8100

TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC 8150
 CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCT 8200
 TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG 8250
 5 CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG 8300
 CGGTGCTACA GAGTTCTGA AGTGGTGGCC TAACTACGGC TACACTAGAA 8350
 GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAAA 8400
 AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG GTAGCGGTGG 8450
 TTTTTTGTIT TGCAAGCAGC AGATTAACGCG CAGAAAAAAA GGATCTCAAG 8500
 10 AAGATCCTTT GATCTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAAC 8550
 TCACGTTAAG GGATTTGGT CATGAGATTA TCAAAAAGGA TCTTCACCTA 8600
 GATCCTTTTA AATTTAAAT GAAGTTTAA ATCAATCTAA AGTATATATG 8650
 AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC 8700
 TCAGCGATCT GTCTATTTCG TTCATCCATA GTTGCCTGAC TCCCCGTCTG 8750
 GTAGATAACT ACGATACGGG AGGGCTTACC ATCTGGCCCC AGTGTGCAA 8800
 15 TGATACCGCG AGACCCACGC TCACCCGCTC CAGATTTATC AGCAATAAAC 8850
 CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT GTTCTGCAA CTTTATCCGC 8900
 CTCCATCCAG TCTATTAAATT GTTGCCTGGGA AGCTAGAGTA AGTAGTTCGC 8950
 CAGTTAATAG TTTGCGAAC GTTGGTGCCTA TTGCTACAGG CATCGTGGTG 9000
 TCACGCTCGT CGTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC 9050
 20 AAGGCGAGTT ACATGATCCC CCATGTTGTG CAAAAAAAGCG GTTAGCTCCT 9100
 TCGGTCTCCTC GATCGTTGTC AGAAGTAAGT TGGCCGCAGT GTTATCACTC 9150
 ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCATGC CATCCGTAAG 9200
 ATGCTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTC TGAGAATAGT 9250
 GTATGCGCG ACCGAGTTGC TCTTGCCCG CGTCAATACG GGATAATACC 9300
 25 GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC 9350
 GGGGCGAAAAA CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTCGATGT 9400
 AACCCACTCG TGCAACCAAC TGATCTTCAG CATCTTTAC TTTCACCAGC 9450
 GTTTCTGGGT GAGCAAAAC AGGAAGGCAA AATGCCGCAA AAAAGGGAAT 9500
 AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCCTT TTCAATATT 9550
 30 ATTGAAGCAT TTATCAGGGT TATTGTCTCA TGAGCGGATA CATATTGAA 9600
 TGTATTTAGA AAAATAAACAA AATAGGGGTT CCGCGCACAT TTCCCCGAAA 9650
 AGTGCCAC 9658

35 SEQ ID NO:5 (spacer)
 (GPGG)_x

40 SEQ ID NO:6 (spacer)
 GPGGGPGGGPGG

45 SEQ ID NO:7 (spacer)
 GGGGSGGGGGSGGGGSGGGGS

50 SEQ ID NO:8 (spacer)
 GGGGSGGGGGSGGGGSGGGGS
 SEQ ID NO:9 (enterokinase cleavage site)
 DDDDK

55 SEQ ID NO:10 (altered transposase Hef forward primer)
 ATCTCGAGACCATGTGTAACTTGATATTTACATGATTCTCTTAC

60 SEQ ID NO:11 (altered transposase Her reverse primer)
 GATTGATCATTATCATAATTCCCCAAAGCGTAACC

SEQ ID NO:12 (Xho I restriction site)
CTCGAG

5 SEQ ID NO:13 (modified Kozak sequence)
ACCATG

10 SEQ ID NO:14 (Bcl I restriction site)
TGATCA

15 SEQ ID NO:15 (CMVf-NgoM IV primer)
TTGCCGGCATCAGATTGGCTAT

20 SEQ ID NO:16 (Syn-polyAr-BstE II primer)
AGAGGTCACCGGGTCAATTCTTCAGCACCTGGTA

25 SEQ ID NO:17 (vitellogenin promoter)
TGAATGTGTT CTTGTGTTAT CAATATAAAT CACAGTTAGT GATGAAGTTG GCTGCAAGCC
TGCATCAGTT CAGCTACTTG GCTGCATTT GTATTTGGTT CTGTAGGAAA TGCAAAAGGT
TCTAGGCTGA CCTGCACCTTC TATCCCTCTT GCCTTAUTGC TGAGAATCTC TGCAAGGTTT
AATTGTTCAC ATTTGCTCC CATTACTTT GGAAGATAAA ATATTTACAG AATGCTTATG
AAACCTTTGT TCATTTAAAAT ATTCTCTGG TCAGCGTGC CGGAGCTGAA AGAACACATT
GATCCCGTGA TTTCATAAAAATACATATGTT CCATATATTG TTTCTCAGTA GCCTCTAAA
TCATGTGCGT TGGTGCACAT ATGAATACAT GAATAGCAA GGTTTATCTG GATTACGCTC
30 TGGCCTGCAAGAATGGCAGTAAACCAAAGCTGAGGGAAAGA GGGAGAGTAT AGTCAATGTA
GATTATACTG ATTGCTGATT GGGTTATTAT CAGCTAGATA ACAACTTGGG TCAGGTGCCA
GGTCAACATA ACCTGGGCAA AACCAGTCTC ATCTGTGGCA GGACCATGTA CCAGCAGCCA
GCCGTGACCC AATCTAGGAA AGCAAGTAGC ACATCAATT TAAATTTATT GTAAATGCCG
TAGTGAAGT GTTTTACTGT GATACATTGA AACTTCTGGT CAATCAGAAA AAGGTTTTT
35 ATCAGAGATGCCAAGGTATT ATTGATTCTTCTTCTTCTCGC CGTGAAGAGA ATTTATGATT
GCAAAAGAG GAGTGTACAT AAAAACTGAGGAAATTCAAG CAGAAACAG
CCACGTGTTCTGAACATTC TTCCATTTAATGCTCACCAGTGCCTGGCAGA GGCCTATTCA
CCTTCGCT

40 SEQ ID NO:18 (vitellogenin targeting sequence)
ATGAGGGGGATCATACTGGCATTAGTGCCTCACCTGTAGGCAGCCAGAAGTTGACATTGGT

45 SEQ ID NO:19 (p146 protein)
KYKKALKKLAKLL

50 SEQ ID NO:20 (p146 coding sequence)
AAATACAAAAAAGCACTGAAAAAACTGGCAAAACTGCTG

55 SEQ ID NO:21 (pro-insulin sequence)
TTTGTGAACCAACACACTGTGCGGCTCACACCTGGTGGAAAGCTCTACCTAGTGTGCGGGGAAACGAGGC
TTCTTCTACACACCCAAAGACCCGCGGGAGGCAGAGGACCTGCAGGTGGGGCAGGTGGAGCTGGCGGG
GCCCTGGTGCAGGCAGCCTGCAGCCCTGGCCCTGGAGGGTCCCTGCAGAAGCGTGGCATTGTGGAA
CAATGCTGTACCAAGCATCTGCTCCCTACCAAGCTGGAGAAACTCTGCAACTAG

60 SEQ ID NO:22 (TAG sequence)

Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp
 Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Thr Thr Cys Ile Leu Lys Gly Ser Cys
 Gly Trp Ile Gly Leu Leu Asp Asp Asp Asp Lys

5 SEQ ID NO:23 (gp41 epitope)
 Ala Thr Thr Cys Ile Leu Lys Gly Ser Cys Gly Trp Ile Gly Leu Leu

10 SEQ ID NO:24 (polynucleotide sequence encoding gp41 epitope)
 Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Thr Thr Cys Ile Leu Lys Gly
 Ser Cys Gly Trp Ile Gly Leu Leu Asp Asp Asp Asp Lys

15 SEQ ID NO:25 (repeat domain in TAG spacer sequence)
 Pro Ala Asp Asp Ala

20 SEQ ID NO:26 (TAG spacer sequence)
 Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp
 Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp

25 SEQ ID NO:27 (Vit pro/Vit targ/TAG/pro-insulin/synthetic polyA)
 TGAATGTGTT CTTGTGTAT CAATATAAT CACAGTTAGT GATGAAGTTG GCTGCAAGCC
 TGCATCAGTT CAGCTACTTG GCTGCATT TT GTATTTGGTT CTGTAGGAAA TGCAAAAGGT
 TCTAGGCTGA CCTGCACCTTC TATCCTCTT GCCTACTGC TGAGAATCTC TGCAGGTTTT
 AATTGTTAC CATTGCTCC CATTACTTT GGAAGATAAA ATATTACAG AATGCTTATG
 AAACCTTTGT TCATTTAAA ATATTCCTGG TCAGCGTGAC CGGAGCTGAA AGAACACATT
 GATCCCGTGA TTTCAATAAA TACATATGTT CCATATATTG TTTCTCAGTA GCCTCTAAA
 TCATGTGCGT TGGTGCACAT ATGAATACAT GAATAGCAAA GGTTTATCTG GATTACGCTC
 30 TGGCCTGCAG GAATGGCCTT AAACCAAAGC TGAGGGAAAGA GGGAGAGTAT AGTCAATGTA
 GATTATACTG ATTGCTGATT GGGTTATTAT CAGCTAGATA ACAACTTGGG TCAGGTGCCA
 GGTCAACATA ACCTGGGCAA AACCAAGTCTC ATCTGTGGCA GGACCATGTA CCAGCAGGCCA
 GCGGTGACCC AATCTAGGAA AGCAAGTAGC ACATCAATT TAAATTATT GTAAATGCCG
 TAGTAGAACT GTTTACTGT GATACATTGA AACTCTGGT CAATCAGAAA AAGTTTTTT
 35 ATCAGAGATG CCAAGGTATT ATTTGATT TTCTTATTCGC CGTGAAGAGA ATTTATGATT
 GCAAAAGAG GAGTGTATTAC ATAAACTGAT AAAAAACTTG AGGAATTCAAG CAGAAAACAG
 CCACGTGTTCTG CTGAACATTC TTCCATTTA GTCTCACCAT GCCTGGCAGA GCCCTATTCA
 CCTTCGCTAT GAGGGGGATC ATACTGGCAT TAGTGCTCAC CCTTGCTAGGC AGCCAGAAAGT
 TTGACATTGG TAGACTGAGA ATGGCAAGAA GAATGAGAAGA TGGTTTGTG AACCAACACC
 40 TGTGCGCTCA CACCTGGTGG AAGCTCTCTA CCTAGTGTGCG GGGAAACGAGG CTTCTTCTAC
 ACACCCAAGA CCCGCCGGGA GGCAGAGGAC CTGCAGGTGGG GCAGGTGGAG CTGGGGGGGG
 GCCCTGGTGC AGGCAGCCTG CAGCCCTGG CCCTGGAGGGG TCCCTGCAGA AGCGTGGCAT
 TGTGGAACAA TGCTGTACCA GCATCTGCTC CCTCTACCAAGC TGGAGAACTA CTGCAACTAG
 GGCAGCTGATCCAGATCACTTCTGGCTAATAAAAGATCAGAGCTCTAGAGATCTGTGTGGTTTT
 45 TGTGGATCTGCTGTGCTTCTAGTTGCGCAGCCATCTGTGTTGCCCTCCCCGTGCCTTCTTGACC
 CTGGAAGGTGCCACTCCCCTGCTTCTAATAAAATGAGGAATTGCACTGCATTGCTGAGTAGG
 TGTCATTCTATTCTGGGGGTGGGTGGGCAGCACAGCAAGGGGAGGATTGGGAAGACAATAGCAGG
 CATGCTGGGATGCGGTGGCTATGGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC
 50 TCTCGGTACCTCTCTC
 55 SEQ ID NO:28 (synthetic polyA sequence)
 GGCAGCTGATCCAGATCACTTCTGGCTAATAAAAGATCAGAGCTCTAGAGATCTGTGTGGTTTT
 TGTGGATCTGCTGTGCTTCTAGTTGCGCAGCCATCTGTGTTGCCCTCCCCGTGCCTTCTTGACC
 CTGGAAGGTGCCACTCCCCTGCTTCTAATAAAATGAGGAATTGCACTGCATTGCTGAGTAGG
 TGTCATTCTATTCTGGGGGTGGGTGGGCAGCACAGCAAGGGGAGGATTGGGAAGACAATAGCAGG
 CATGCTGGGATGCGGTGGCTATGGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC
 TCTCGGTACCTCTCTC

SEQ ID NO:29 (pTnMod(Oval/ENT tag/P146/PA) - Chicken)

5	CTGACGCGCC	CTGTAGCGGC	GCATTAAGCG	CGGCGGGTGT	GGTGGTTACG	50
	CGCAGCGTGA	CCGCTACACT	TGCCAGCGCC	CTAGCGCCCG	CTCCTTCGC	100
	TTTCTTCCCT	TCCTTCTCG	CCACGTCGC	CGGCATCAGA	TTGGCTATTG	150
	GCCATTGCAT	ACGTGATAC	CATATCATAA	TATGTACATT	TATATTGGCT	200
	CATGTCAAC	ATTACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	250
10	TAGTAATCAA	TTACGGGTC	ATTAGTCAT	AGCCCCATATA	TGGAGTTCG	300
	CGTTACATATA	CTTACGGTAA	ATGGCCGCC	TGGCTGACCG	CCCAACGACC	350
	CCCGCCCAT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	400
	GGGACTTTCC	ATTGACGTCA	ATGGTGGAG	TATTTACGGT	AAACTGCCA	450
	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAAGTACGCC	CCTATTGACG	500
15	TCAATGACGG	TAATGGCCC	GCCTGGCATT	ATGCCAGTA	CATGACCTTA	550
	TGGGACTTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTC	TCGCTATTAC	600
	CATGGTGATG	CGGTTTGGC	AGTACATCAA	TGGGCGTGG	TAGCGGTTTG	650
	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TGACGCTCAA	TGGGAGTTTG	700
	TTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGTCGTA	ACAACCTCCG	750
20	CCCATGGAGC	CAAATGGCG	GTAGGGTGT	ACGGTGGAG	GTCTATATAA	800
	GCAGAGCTCG	TTTACTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACGC	850
	TGTTTGACCC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGCCG	900
	GGAACGGTGC	ATTGGAACGC	GGATTCCCCG	TGCAAGAGT	GACGTAAGTA	950
	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA	1000
25	CTGTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTTCCTTAT	GCTATAGGTG	1050
	ATGGTATAGC	TTAGCCTATA	GGTGTGGTT	ATTGACCAT	ATTGACCACT	1100
	CCCCTATTGG	TGACGATACT	TTCCATTACT	AATCCAAAC	ATGGCTCTTT	1150
	GCCACAACTA	TCTCTATGG	CTATATGCCA	ATACTCTGTC	CTTCAGAGAC	1200
	TGACACGGAC	TCTGTATT	TACAGGATGG	GGTCCCATT	ATTATTTACA	1250
30	AATTACACATA	TACAACAAACG	CCGTCACCCCG	TGCCCCGAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAATCTCGG	GTACGTGTC	CGGACATGGG	1350
	CTCTTCTCCG	GTAGCGGCCG	AGCTTCCACA	TCCGAGCCCT	GGTCCCAGTC	1400
	CTCCAGCGGC	TCATGGTCGC	TCGGCAGCTC	CTTGCTCCTA	ACAGTGGAGG	1450
	CCAGACTTAG	GCACAGCACA	ATGCCACCA	CCACCACTGT	GCCGCAACAG	1500
35	GCCGTGGCG	TAGGGTATGT	GTCTGAAAT	GAGCGTGGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGAA	GACTTAAGGC	AGCGGCAGAA	GAAGATGCA	1600
	GCAGCTGAGT	TGTTGTATT	TGATAAGAGT	CGAGGGTAAC	TCCCCTGTCG	1650
	GTGCTGTTAA	CGGTGGAGGG	CAGTGTAGTC	TGAGCAGTAC	TCGTTGCTGC	1700
	CGCGCGCGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCTTT	1750
40	CCATGGGTCT	TTTCTGCACT	CACCTCGGA	CCATGTGTA	ACTTGATATT	1800
	TTACATGATT	CTCTTACCA	ATTCTGCC	GAATTACACT	AAAAACGACT	1850
	CAACAGCTTA	ACGTTGGCT	GCCACGATT	ACTTGACTGT	AAAACCTCTA	1900
	CTCTTACCGA	ACTTGGCCGT	AACCTGCCA	CCAAAGCGAG	AACAAAACAT	1950
	ACACATCAAAC	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAAGA	2000
45	GCGACTCGCT	GTATACCGTT	GGCATGCTAG	CTTATATCTGT	TCGGGAATAC	2050
	GATGCCATT	GTACTTGTG	ACTGGCTCT	TATTCGTGAG	AAAAAACGAC	2100
	TTATGGTATT	GCGAGCTTCA	GTGCACTAC	ACGGTCGTT	TGTTACTCTT	2150
	TATGAGAAAG	CGTTCCCGCT	TTCAAGCAGA	TGTTCAAGA	AAGCTCATGA	2200
	CCAATTCTCA	GGCGACCTTG	CGACGATTCT	ACCGAGTAAC	ACCACACCGC	2250
50	TCATTGTCAG	TGATGCTGG	TTTAAAGTGC	CATGGTATAA	ATCCGTTGAG	2300
	AAGCTGGTT	GGTACTGGTT	AAGTCAGTA	AGAGGAAAAG	TACAATATGC	2350
	AGACCTAGGA	GCGGAAAAGT	GGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
	CATCTAGTC	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	AAAAAGCAAT	2450
	CCAATCTCAT	GCCAAATCT	ATTGTATAAA	TCTCGCTCTA	AAGGCCGAAA	2500
55	AAATCAGCGC	TCGACACCGA	CTCATGTCA	CCACCGCTCA	CCTAAACATCT	2550
	ACTCAGCGTC	GGCAAAGGAG	CCATGGGTC	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTGCAA	CACCCAAACA	ACTTGTAAAT	ATCTATTGCA	AGCGAATGCA	2650
	GATTGAAGAA	ACCTTCCGAG	ACTTGAAGAAAG	TCTCGCTCTA	GGACTAGGGCC	2700
	TACGCCATAG	CCGAACGAGC	AGCTCAGAGC	GTTTTGATAT	CATGCTGCTA	2750
60	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	CTTGGGGCG	TTCATGCTCA	2800
	GAAACAAGGT	TGGGACAAGC	ACTTCCAGGC	TAACACAGTC	AGAAATCGAA	2850
	ACGTACTCTC	AACAGTTCGC	TTAGGCATGG	AAAGTTTGC	GCATTCTGGC	2900

TACACAATAA CAAGGGAAAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTTATTC ACACATGGTT ACGCTTGGG GAAATTATGA TAATGATCCA 3000
 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 5 TTTTTGTTGG ATCTGCTGTG CCTCTAGTT GCCAGCCATC TGTTGTTGC 3100
 CCCTCCCCCG TGCCCTCCT GACCTGGAA GGTGCCACTC CCACTGTCC 3150
 TTCCTAATAA AATGAGGAAA TTGATCGCA TTGCTCTGAGT AGGTGTCATT 3200
 CTATTCTGGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGG GGATTGGGAA 3250
 GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 10 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCGGTAC TGCTGAAGAA 3400
 TTGACCCGGT GACCAAAGGT GCCTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCCTGTTA TAAGCAGCAA TTAATTATGA TTGATGCC 3500
 CATCACAAACA AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTGAACA TTATCTTGAT TATATTATTG ATAATAATAA AAACCTTATC 3600
 15 CCTATCCAAG AAGTGTGCC TATCATTGGT TGAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGGTAAC CGCCATTGT CAGCAAATTG 3700
 ATCCAAGAGA ACCAACTTAA AGCTTCTCTG ACAGGAATGTT AATTCTCGTT 3750
 GACCCCTGAGC ACTGATGAAT CCCCTAATGA TTGGTAAAGG AATCATTAAAG 3800
 20 TTAAGGTGGA TACACATCTT GTCATATGAT CCCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCCAAGG CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGGAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGCGC AATTAACCCCT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTAGTGGA TCCCCCGGGG 4050
 AGGTCAAGAAT GGTTCCTTA CTGTTGTCA ATTCTATTAT TTCAATACAG 4100
 25 ACAATAGCT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG 4150
 TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTCACAA TTCCCTCTGTC 4200
 ATCTGCCAGG CCATTAAGTT ATTCAATGGAA GATCTTTGAG GAACACTGCA 4250
 AGTTCATATC ATAACACAT GTTAAATTGA GTATGTTTT GCATTGTATG 4300
 GAGCTATGTT TTGCTGTATC CTCAAGAAAAA AAGTTGTTA TAAAGCATTC 4350
 30 ACACCCATAA AAAGATAGAT TAAATATTC CAGCTATAGG AAAGAAAGTG 4400
 CGTCTGCTCT TCACTCTAGT CTCAGTTGGC TCCTTCACAT GCATGCTTCT 4450
 TTATTTCTCC TATTTGTCA AGAAAATAAT AGGTCACTGC TTGTTCTCAC 4500
 TTATGCTCTG CCTAGCATGG CTCAGATGCA CGTTGAGAT ACAAGAAGGA 4550
 TCAAATGAAA CAGACTCTG GTCTGTTACT ACAACCATAG TAATAAGCAC 4600
 35 ACTAACTAAT AATTGCTAAT TATGTTTTCC ATCTCTAAAGG TTCCACATT 4650
 TTTCTGTTT CTTAAAGATC CCATTATCTG GTTGTAACTG AAGCTCAATG 4700
 GAACATGAGC AATATTTCC AGTCTCTCT CCCATCCAAC AGTCTGTATG 4750
 GATTAGCAGA ACAGGGAGAA AACACATTGT TACCCAGAAAT TAAAACCTAA 4800
 TATTGCTCT CCATTCAATC CAAAATGGAC CTATTGAAAC TAAAATCTAA 4850
 40 CCCAATCCCA TAAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG 4900
 GGAACCTGTG GGTGGTCAC AATTCAAGCT ATATATTCCC CAGGGCTCAG 4950
 CGGATCCATCG GGCTCCATCG GCGCAGCAAG CATGGAAATT TGTTTGATG 5000
 TATTCAAGGA GCTCAAAGTC CACCATGCCA ATGAGAACAT CTTCTACTGC 5050
 CCCATTGCCA TCATGTCAGC TCTAGGCCATG GTATACCTGG GTGCAAAAGA 5100
 45 CAGCACCAGG ACACAGATAA ATAAGGTTGT TCGCTTTGAT AAACCTCCAG 5150
 GATTGGAGA CAGTATTGAA GCTCAGTGTG GCACATCTGT AAACGTTCAC 5200
 TCTTCACCTA GAGACATCTC CAACCAAATC ACCAACCAAATGATGTTA 5250
 TTGCTTCAGC CTTGCCAGTA GACTTTATGC TGAAGAGAGA TACCCAATCC 5300
 TGCCAGAATA CTTGCAGTGT GTGAAGGAAC TGATAGAGG AGGCTTGAA 5350
 50 CCTATCAACT TTCAAACAGC TGCAAGATCAA GCCAGAGAGC TCATCAATTG 5400
 CTGGGTAGAA AGTCAGACAA ATGCAATTAT CAGAAATGTC CTTCAGCCAA 5450
 GCTCCGTGGA TTCTCAAACG GCAATGGTTC TGTTAATGC CATTGCTCTC 5500
 AAAGGACTGT GGGAGAAAATTTAAGGAT GAAGACACAC AAGCAATGCC 5550
 TTTCAGAGTG ACTGAGCAAG AAAGCAAACCC TGTGAGATG ATGTACCAAGA 5600
 55 TTGGTTTATT TAGAGTGGCA TCAATGGCTT CTGAGAAAAT GAAGATCTG 5650
 GAGCTTCCAT TTGCCAGTGG GACAATGAGC ATGTTGGTGC TGTTGCC 5700
 TGAAGTCTCA GGCCTTGAGC AGCTTGAGAG TATAATCAAC TTTGAAAAC 5750
 TGACTGAATG GACCAGTTCT AATGTTATGG AAGAGAGGAA GATCAAAGTG 5800
 TACTTACCTC GCATGAAGAT GGAGGAAAAA TACAACCTCA CATCTGTCTT 5850
 60 AATGGCTATG GGCATTACTG ACGTGTTTAG CTCTTCAGCC AATCTGTCTG 5900
 GCATCTCCTC AGCAGAGAGC CTGAAGATAT CTCAAGCTGT CCATGCCAGCA 5950

CATGCAGAAA TCAATGAAGC AGGCAGAGAG GTGGTAGGGT CAGCAGAGGC 6000
 TGGAGTGGAT GCTGCAAGCG TCTCTGAAGA ATTAGGGCT GACCATCCAT 6050
 TCCTCTTCTG TATCAAGCAC ATCGCAACCA ACGCCGTTCT CTTCTTGGC 6100
 5 AGATGTGTTT CCCCTCCGCG GCCAGCAGAT GACGCACCAG CAGATGACGC 6150
 ACCAGCAGAT GACGCACCAG CAGATGACGC ACCAGCAGAT GACGCACCAG 6200
 CAGATGACGC AACAAACATGT ATCCCTGAAAG GCTCTTGTTG CTGGATGGC 6250
 CTGCTGGATG ACGATGACAA AAAATACAAA AAAGCACTGA AAAACTGGC 6300
 AAAACTGCTG TAATGAGGGC GCCTGGATCC AGATCACTTC TGGCTAATAA 6350
 10 AAGATCAGAG CTCTAGAGAT CTGTTGTTG GTTTTTGTTG GATCTGCTGT 6400
 GCCTCTAGT TGCCAGCAT CTGTTGTTG CCCCTCCCCC GTGCTTCTC 6450
 TGACCCCTGGA AGGTGCCACT CCCACTGTCC TTTCTAATAA AAATGAGGAA 6500
 ATTGCATCGC ATTGTCTGAG TAGGTGTATC TCTATTCTGG GGGGTGGGGT 6550
 GGGGCAGCAC AGCAAGGGG AGGATTGGGA AGACAATAGC AGGCATGCTG 6600
 GGGATGCGGT GGGCTCTATG GGTACCTCTC TCTCTCTCTC TCTCTCTC 6650
 15 TCTCTCTCTC TCTCTCGTA CCTCTCTCGA GGGGGGGCCC GGTACCCAAT 6700
 TCGCCCTATA GTGAGTCGTA TTACGGCGC TCACTGGCCG TCGTTTACA 6750
 ACGTCGTGAC TGGGAAACCC CTGGCGTTAC CCAACTTAAT CGCCTTGCG 6800
 CACATCCCCC TTTGCCAGC TGCGTAATAA GCGAAGGGC CGGCACCGAT 6850
 CGCCCTTCCC AACAGTTGCG CAGCCTGAAT GGCAGATGGA AATTGTAAGC 6900
 20 GTTAATATTT TGTTAAAATT CGCGTTAAAT TTTTGTAAAT TCAGCTCATT 6950
 TTTTAACCAA TAGGCCGAAA TCGGCAAAT CCCTTATAAA TCAAAAGAAT 7000
 AGACCGAGAT AGGGTTGAGT GTTGTCCAG TTGGAACAA GAGTCCACTA 7050
 TTAAAGAACG TGGACTCCAA CGTCAAAGGG CGAAAACCG TCTATCAGGG 7100
 CGATGGCCCA CTACTCCGGG ATCATATGAC AAGATGTGTA TCCACCTTAA 7150
 25 CTTAATGATT TTTACCAAAA TCATTAGGG ATTCACTAGT GCTCAGGGTC 7200
 AACAGAGAATT AACATTCCGT CAGGAAAGCT TATGATGATG ATGTGCTTAA 7250
 AAACCTTACTC AATGGCTGGT TATGCATATC GCAATACATG CGAAAAAACCT 7300
 AAAAGAGCTT GCCGATAAAA AAGGCCAATT TATTGCTATT TACCGCGGCT 7350
 TTTTATTGAG CTTGAAAGAT AAATAAAATA GATAGGTTTT ATTGAGACT 7400
 30 AAATCTTCTT TATCGTAAAA AATGCCCTCT TGTTGTTATCA AGAGGGTCAT 7450
 TATATTCGCG GGAATAACAT CATTGGTGA CGAAATAACT AAGCACTTGT 7500
 CTCCCTGTTA CTCCCTGAG CTTGAGGGT TAACATGAAG GTCATCGATA 7550
 GCAGGATAAT AATACAGTAA AACGCTAAC CAAATAATCCA AATCCAGCCA 7600
 TCCCAAATTG GTACTGAATG ATTAAATAA ACAGCAAAAC GAAATGGGCC 7650
 35 AATAAACACCG GTTGCATGG TAAGGCTCAC CAATAATCCC TGAAAGCAC 7700
 CTTGCTGATG ACTCTTTGTT TGATAGAAC TCACTCCCTG TAATGCAGGT 7750
 AAAGCGATCC CACCAACCGC CAATAAAATT AAAACAGGGG AACTAACCA 7800
 ACCTTCAGAT ATAAACGCTA AAAAGGCAA TGCACTACTA TCTGCAATAA 7850
 ATCCGAGCAG TACTGCCGTT TTTCGCCCC ATTAGTGGC TATTCTTCTC 7900
 40 GCCACAAAGG CTTGGAATAC TGAGTGTAAA AGACCAAGAC CCGCTAATGA 7950
 AAAGCCAACC ATCATGCTAT TCCATCCAAA ACGATTTTCG GTAAATAGCA 8000
 CCCACACCGT TGCGGGATT TGGCTATCA ATTGCGCTGA AAAATAAAATA 8050
 ATCAACAAAATG TGGCATCGTT TTAATAAAAG TGATGTATAC CGAATTTCAGC 8100
 TTTGTTTCCC TTAGTGGAGG GTTAATTGCG CGCTTGGCGT AATCATGGTC 8150
 45 ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAAACA 8200
 TACGAGCCGG AAGCATAAAG TGTAAGCCT GGGGTGCTA ATGAGTGAGC 8250
 TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGTTTCC AGTCGGGAAA 8300
 CCTGTCGTGC CAGCTGCATT AATGAATCGG CCAACGCGC GGGAGAGGGC 8350
 GTTGTGCTAT TGGGCGCTCT TCCGCTTCTC CGCTCACTGA CTCGCTGGC 8400
 50 TCGGTCGTTC GGCTCGGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAAT 8450
 ACGGTTATCC ACAGAACATCG GGGATAACGC AGGAAAGAAC ATGTGAGCAA 8500
 AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT 8550
 TTCCATAGGC TCCGCCCCC TGACGAGCAT CACAAAATC GACGCTCAAG 8600
 TCAGAGGTGG CGAAACCGA CAGGACTAA AGATACCCAG GCGTTCCCCC 8650
 55 CTGGAAGCTC CCTCGTGCCTC TCTCTCTTCC CGACCCCTGCC GCTTACCGGA 8700
 TACCTGTCCG CCTTTCTCTC TTGCGGAAGC GTGGCGCTTT CTCATAGCTC 8750
 ACGCTGTAGG TATCTCAGTT CGGTGTTAGGT CGTTCGCTCC AAGCTGGGCT 8800
 GTGTGCACGA ACCCCCCCTT CAGCCGACCC GCTGCGCCTT ATCCGTAAC 8850
 TATCGTCTTG AGTCCAACCC GGTAAAGACAC GACTTATCGC CACTGGCAGC 8900
 60 AGCCACTGGT AACAGGATTA GCAGAGGGAG GTATGTAGGC GGTGCTACAG 8950
 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAG GACAGTATT 9000

GGTATCTGCG CTCTGCTGAA GCCAGTTACG TTTCGGAAAAA GAGTTGGTAG 9050
 CTCTTGATCC GGCACACAAA CCACCGCTGG TAGCGGTGGT TTTTTGTTT 9100
 GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG 9150
 ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC CACGTTAAGG 9200
 5 GATTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTAA 9250
 ATTAAAAATG AAGTTTAA TCAATCTAAA GTATATATGA GTAAACTTGG 9300
 TCTGACAGTT ACCAATGCTT AATCACTGAG GCACCTATCT CAGCGATCTG 9350
 TCTATTTCGT TCATCCATAG TTGCGTACT CCCCGTCGTG TAGATAACTA 9400
 CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA 9450
 10 GACCCACGCT CACCGCTCC AGATTATCA GCAAAATAAC AGCCAGCCGG 9500
 AAGGGCCGAG CGCAGAAGTG GTCTCGAAC TTTATCCGCC TCCATCCAGT 9550
 CTATTAATTG TTGCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAAAGT 9600
 TTGCGCAACC TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC 9650
 GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA 9700
 15 CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCCTCCG 9750
 ATCGTTGTCA GAAGTAAGTT GGCGCAGTG TTATCACTCA TGGTTATGGC 9800
 AGCACTGCT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTCTG 9850
 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGGCGA 9900
 CCGAGTTGCT CTTGCCCCGC GTCAATACGG GATAATACGG CGCCACATAG 9950
 20 CAGAACTTAA AAGTGTCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC 10000
 TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT 10050
 GCACCCAACG GATCTTCAGC ATCTTTACT TTCAACCGCG TTTCTGGGTG 10100
 AGCAAAAACA GGAAGGCAAA ATGCCGAAA AAAGGGAATA AGGGCGACAC 10150
 GGAAATGTTG AATACTCATA CTCTTCCTT TTCAATATTA TTGAAGCATT 10200
 25 TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTGAA GTATTTAGAA 10250
 AAATAAACAA ATAGGGGTTG CGCGCACATT TCCCCGAAAAA GTGCCAC 10297

SEQ ID NO:30 (pTnMod(Oval/ENT tag/P146/PA) - QUAIL)

30 CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCCTTCGC 100
 TTTCTTCCCT TCCTTTCTCG CCACGTTCGC CGGCATCAGA TTGGCTATTG 150
 GCCATTGCT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
 35 CATGTCACAC ATTACGCCA TTGTCACATT GATTATTGAC TAGTTATTAA 250
 TAGTAACTCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTCCG 300
 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGCGTGACCG CCCAACGACC 350
 CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AACTGCCA 450
 40 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG 500
 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA 550
 TGGGACTTTTC CTACTTGCA GTACATCTAC GTATTAGTC TCGCTATTAC 600
 CATGGTGTAG CGGTTTTGGC AGTACATCAA TGCGCTGGA TAGCGGTTTG 650
 ACTCACGGGG ATTTCACAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG 700
 45 TTTGGCACC AAAATCAACG GGACTTTCCA AATGTCGTA ACAACTCCGC 750
 CCCATTGAGC CAAATGGGGC GTAGCGTGT ACGGTGGAG GTCTATATAA 800
 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
 TGTTTGACC TCCATAGAAG ACACCGGGAC CGATCAGGCC TCCGCGGCCG 900
 GGAACGGTGC ATTGGAACGC GGATCCCCG TGCCAAGAGT GACGTAAGTA 950
 50 CCGCCTATAG ACTCTATAGG CACACCCCTT TGCGCTTTAT GCATGCTATA 1000
 CTGTTTTGCG CTTGGGGCCT ATACACCCCCC GCTTCCTTTAT GCTATAGGTG 1050
 ATGGTATAGC TTAGCCTATA GGTGTGGTT ATTGACCTT ATTGACCACT 1100
 CCCCTATTG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
 GCCACAACTA TCTTCTATGG CTATATGCC ATACTCTGTC CTTCAGAGAC 1200
 55 TGACACGGAC TCTGTATTT TACAGGATGG GGTCCCATT ATTATTTACA 1250
 AATTACACATA TACAACAACG CCGTCCCCG TGCGCGCAGT TTTTATTAAA 1300
 CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTT CGGACATGGG 1350
 CTCTTCTCCG GTAGCGGGGG AGCTTCCACA TCCGAGCCCT GGTCCCAGTC 1400
 CTCCAGCGGC TCATGGTCGC TCAGCAGCTC CTTGCTCCTA ACAGTGGAGG 1450
 60 CCAGACTTAG GCACAGCACAA ATGCCACCA CCACCACTGT GCCGCACAAG 1500
 GCCGTGGCGG TAGGGTATGT GTCTGAAAAT GAGCGTGGAG ATTGGGCTCG 1550

CACGGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG 1600
 GCAGCTGAGT TGTGTATTG TGATAAGAGT CAGAGGTAAC TCCCCTTGC 1650
 GTGCTTAA CGGTGGAGGG CAGTAGTC TGAGCAGTAC TCGTTGCTGC 1700
 CGCGCGCAGC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTT 1750
 5 CCATGGGTCT TTTCTGCAGT CACCGTCGGA CCATGTGTGA ACTTGATATT 1800
 TTACATGATT CTCTTTACCA ATTCCTCCCC GAATTACACT TAAAACGACT 1850
 CAACAGCTTA ACCTTGGCTT GCCACCGATT ACTTGACTGT AAAACTCTCA 1900
 CTCTTACCGA ACTTGGCGT AACCTGCCA CCAAAGCGAG AACAAAACAT 1950
 AACATCAAAC GAATCGACCG ATTGTTAGGT AATCGTCACC TCCACAAAGA 2000
 10 GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAAATAC 2050
 GATGCCATT GTACTTGTG ACTGTCGTGA TATTGTCGAG CAAAAACGAC 2100
 TTATGGTATT GCGAGCTCA GTCGCACTAC ACGGTCGTT TGTTACTCTT 2150
 TATGAGAAAG CGTTCCCGCT TTCAGAGCAA TGGTCAAAGA AAGCTCATGA 2200
 CCAATTCTCA GCCGACCTTG CGAGCATTCT ACCGAGTAAC ACCACACCGC 2250
 15 TCATTGTCAG TGATGCTGGC TTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
 AGACTGGTT GGTACTGGTT AAGTCGAGTA AGAGGAAAG TACAATATGC 2350
 AGACCTAGGA GCGAAACCTAT CAGCAACTTA CATGATATGT 2400
 CATCTAGTC CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAAGCAAT 2450
 CCAATCTCAT GCACAAATTCT ATTGTTAAA TCTCGCTCTA AAGGCCGAAA 2500
 20 AAATCAGCGC TCGACACCGA CTCAATGTC CAACCCGTC CCTAAAATCT 2550
 ACTCAGCGTC GGCAAAGGAG CCATGGGTT TAGCAACTAA CTTACCTGTT 2600
 GAAATTGAA CACCCAAACA ACTTGTAAAT ATCTATTGCA AGCGAATGCA 2650
 GATTGAAGAA ACCTTCCGAG ACTTGAAAG TCCCGCCTAC GGACTAGGCC 2700
 TAGGCCATAG CGAACGAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750
 25 ATCGCCCTGA TGCTTCAACT AACATGTTGG CTGGGGGGC TTGATGCTCA 2800
 GAAACAAGGT TGGGACAAGC ACTTCCAGGC TAACACAGTC AGAAATCGAA 2850
 ACGTACTCTC AACAGTTCGTC TTAGGCATGG AAGTTTTGCG GCATTCTGGC 2900
 TACACAATAA CAAGGGAAAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTTATTC ACACATGGTT ACGCTTGGG GAAATTATGA TAATGATC 3000
 30 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 TTTTTGTTG ATCTGCTGT CCTCTAGTT GCCAGCCATC TGTTGTTGC 3100
 CCCTCCCCCG TGCTTCTCTT GACCTGGAA GGTGCCACTC CCACTGTCT 3150
 TTCTCTCTTAA AATGAGGAAA TTGATCGCA TTGTCGTACT AGGTGTCATT 3200
 CTATTGCTGGG GGGTGGGGTG GGGCACACCA GCAAGGGGGG GGATGGGAA 3250
 35 GACAAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCGG TGCTGAAGAA 3400
 TTGACCCGGT GACCAAAAGGT GCCTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCTGTTA TAAGCAGCAA TAAATTATGA TTGATGCCTA 3500
 40 CATCACAAACA AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTTGAACA TTATCTTGAT TATATTATTG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTGGT TGAATGAAC TTGAAAAAAA 3650
 TTAGGCTTGA ATACATTACT GGTAAAGTAA ACGCCATTGT CAGCAAATIG 3700
 ATCCAAGAGA ACCAAACTAA AGCTTCCCTG ACAGGAATGTT AATTCTCGTT 3750
 45 GACCCGTGAGC ACTGATGAAT CCCCTAATGA TTTTGGTAAA AATCATTAAAG 3800
 TTAAGGTGGA TACACATCTT GTCATATGAT CCCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCCAAGG CTTCACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGCGC AATTAACCC CACTAAAGGG AACAAAAGCT 4000
 50 GGAGCTCCAC CGGGTGGCG GCCGCTCTAG AACTAGTGGG TCCCCCGGGG 4050
 AGGTCAAGAT GGTTCTTTA CTGTTGTC ATTCTATTAT TTCAATACAG 4100
 AACAAAAAGCT TCTATAACTG AAATATATTG GCTATTGTAT ATTATGATIG 4150
 TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTCACAA TTCCCTGTC 4200
 ATCTGCCAGG CTGGAGAGTC ATGGAAGATC TCTGAGGAAC ATTGCAAGTT 4250
 55 CATAACCTAA ACTCATTGGA AATTGAGTAT TATTTGCTT TGAATGGAGC 4300
 TATGTTTGCG AGTTCCCTCA GAAGAAAAGC TTGTTATAAA GCGTCTACAC 4350
 CCATCAAAAG ATATATTAA ATATTCAAC TACAGAAAGA TTTTGTCTGC 4400
 TCTTCACTCT GATCTCAGTT GGTTCTTCG CGTACATGCT TCTTTATTG 4450
 CCTATTTGT CAAGAAAATA ATAGGTCAAG TCTGTTCTC ACTTATCTCC 4500
 60 TGCCTAGCAT GGCTTAGATG CACGGTGTAC ATTCAAGAAG GATCAAATGA 4550
 AACAGACTTC TGGTCTGTTA CAACAAACCAT AGTAATAAAC AGACTAACTA 4600

ATAATTGCTA ATTATGTTT CCATCTCTAA GGTTCCCACA TTTTCTGTT 4650
 TTAAGATCCC ATTATCTGGT TGTAACTGAA GCTCAATGGA ACATGAACAG 4700
 TATTCTCAAG TCTTTCTCC AGCAATCCTG ACGGATTAGA AGAACTGGCA 4750
 5 GAAAACACTT TGTTACCCAG AATTAAAAAC TAATATTGTC TCTCCCTTCA 4800
 ATCCAAAATG GACCTATTGA AACTAAAATC TGACCCAATC CCATTAAATT 4850
 ATTTCTATGG CGTCAAAGGT CAAACTTTG AAGGGAACCT GTGGGTGGGT 4900
 CCCAATTCAAG GCTATATATT CCCCAGGGCT CAGCCAGTGG ATCCATGGC 4950
 TCCATCGGTG CAGCAAGCAT GGAATTTTGT TTGATGTAT TCAAGGAGCT 5000
 10 CAAAGTCCAC CATGCCAATG ACAACATGCT CTACTCCCCC TTTGCCATCT 5050
 TGTCACCTCT GGCCATGGTC TTCTAGGTG CAAAAGACAG CACCAGGACC 5100
 CAGATAATAA AGGTTCTCA CTTTGATAAA CTTCAGGAT TCGGAGACAG 5150
 TATTGAAGCT CAGTGTGGCA CATCTGAAA TTGTCACTCT TCACTTAGAG 5200
 ACATACTCAA CCAAATCACC AAACAAAATG ATGCTTATTG TTTCAGCCTT 5250
 GCCAGTAGAC TTTATGCTCA AGAGACATAC ACAGTCGTGC CGGAATACTT 5300
 15 GCAATGTGTG AAGGAACGT ATAGAGGAGG CTTAGAATCC GTCAACTTTC 5350
 AACAGCTGC AGATCAAGCC AGAGGCCTCA TCAATGCCTG GGTAGAAAGT 5400
 CAGACAAACG GAATTATCAG AAACATCCTT CAGCCAAGCT CCCTGGATTC 5450
 TCAAACATGCA ATGGCTCTGG TTAATGCCAT TGCCCTCAAG GGACTGTGGG 5500
 AGAAAAGCATT TAAGGCTAA GACACGCCAA CAATACCTTT CAGAGTGACT 5550
 20 GAGCAAGAAA GCACACCTGT CGAGATGATG TACCAAGATTG GTTCATTAA 5600
 AGTGGCATCA ATGGCTCTG AGAAAATGAA GATCCTGGAG CTTCATTG 5650
 CCAGTGGAAAC AATGAGCATG TTGGTGTGT TGCCGTGATGA TGTCTCAGGC 5700
 CTTGAGCAGC TTGAGAGTAT AATCAGCTT GAAAAGACTGA CTGAATGGAC 5750
 CAGTTCTAGT ATTATGGAAG AGAGGAAGGT CAAAGTGTAC TTACCTCGCA 5800
 25 TGAAGATGGA GGAGAAATAC AACCTCACAT CTCTCTTAAT GGCTATGGGA 5850
 ATTACTGACC TGTTCAGCTC TTCAAGCCAA CTGTCCTGGCA TCTCCTCAGT 5900
 AGGGAGCCTG AAGATATCTC AAGCTGTCA TGACGACAT GCAGAAATCA 5950
 ATGAAGCGGG CAGAGATGTG GTAGGCTCAG CAGAGGCTGG AGTGGATGCT 6000
 ACTGAGAAAT TTAGGCTGA CCATCACATC CTCTCTGTG TCAAGCACAT 6050
 30 CGAAACCAAC GCCATTCTCC TCTTGGCAG ATGTGTTTCT CCCGGCCAG 6100
 CAGATGACGC ACCAGCAGAT GACGCCAACAG CAGATGACGC ACCAGCAGAT 6150
 GACGCACCAG CAGATGACGC ACCAGCAGAT GACGCAACAA CATGTATCCT 6200
 GAAAGGCTCT TGTGGCTGGA TCGGCTGCT GGATGACGAT GACAAAAAAT 6250
 ACAAAAAGC ACTGAAAAAA CTGGCAAAAC TGCTGTAATG AGGGCGCTG 6300
 35 GATCCAGATC ACTTCTGGCT AATAAAAGAT CAGAGCTCTA GAGATCTGTG 6350
 TGTTGGTTT TTGTTGGATCT GCTGTGCTT CTAGTTGCCA GCCATCTGTT 6400
 GTTGGCCCTT CCCCCGTGCC TTCCCTGACC CTGGAAGGTG CCACCTCCAC 6450
 TGTCTTTCC TAATAAAAATG AGGAAATTGCA ATCGCATTGT CTGAGTAGGT 6500
 GTCATTCTAT TCTGGGGGT GGGGTGGGC AGCACAGCAA GGGGGAGGAT 6550
 40 TGGGAAGACA ATAGCAGGCA TGCTGGGAT GCGGTGGGCT CTATGGGTAC 6600
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCTCT 6650
 CTCGAGGGGG GGCCCGGTAC CCAATTGCGC CTATAGTGAG TCGTATTACG 6700
 CGCGCTCACT GGCCGCTGTT TTACAAACGTC GTGACTGGGA AAACCCCTGGC 6750
 GTTACCCAAC TTAATCGCT TGCAGCACAT CCCCCTTCG CCAGCTGGCG 6800
 45 TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC 6850
 TGAATGGCGA ATGGAAATTG TAAGCGTTAA TATTTTGTTA AAATCGCGT 6900
 TAAATTTTG TAAATCAGC TCATTTTTA ACCAATAGGC CGAAATCGGC 6950
 AAAATCCCTT ATAATCAAA AGAATAGACC GAGATAGGGT TGAGTGTGTT 7000
 TCCAGTTGG AACAAGAGTC CACTATTAA GAACGTGGAC TCCAACGTCA 7050
 50 AAGGGCGAAA AACCGTCTAT CAGGGCGATG GCCCACTACT CCGGGATCAT 7100
 ATGACAAGAT GTGTATCCAC CTTAACTTAA TGATTTTTAC CAAAATCATT 7150
 AGGGGATTCA TCACTGCTCA GGGTCAACGA GAATTAACAT TCCGTCAGGA 7200
 AAGCTTATGA TGATGATGTG CTAAAAAAACT TACTCAATGG CTGGTTATGC 7250
 ATATCGCAAT ACATCGAAA AACCTAAAAG AGCTTGGCGA TAAAAAAGGC 7300
 55 CAATTATTTG CTATTACCG CGGCTTTTA TTGAGCTTGA AAGATAATAA 7350
 AAATAGATAG GTTTTATTTG AAGCTAAATC TTCTTTATCG TAAAAAATGC 7400
 CCTCTGGGT TATCAAGAGG GTCAATTAT TTCGCGGAAT AACATCATT 7450
 GGTGACGAAA TAACTAAGCA CTTGTCCTT GTTACTCCC CTGAGCTTGA 7500
 GGGGTTAACAA TGAAGGTCA CGATAGCAGG ATAATAATAC AGTAAAACGC 7550
 60 TAAACCAATA ATCCAAATCC AGCCATCCCA AATTGGTAGT GAATGATTAT 7600
 AAATAACAGC AAACAGTAAT GGGCCAATAA CACCGGTTGC ATTGGTAAGG 7650

5 CTCACCAATA ATCCCTGTAA AGCACCTTGC TGATGACTCT TTGTTTGGAT 7700
 AGACATCACT CCCTGTAATG CAGGTAAGC GATCCCCACCA CCAGCCAATA 7750
 AAATTAAAAC AGGGAAAAGT AACCAACCTT CAGATATAAA CGCTAAAAG 7800
 GCAAATGCAC TACTATCTGC AATAAATCCG AGCAGTACTG CCGTTTTTC 7850
 10 GCCCCATTAA GTGGCTATTC TTCTGCCAC AAAGGCTTGG AATACTGAGT 7900
 GTAAAAGACC AAGACCCGCT AATGAAAAGC CAACCATCAT GCTATTCCAT 7950
 CCAAACGAT TTTCGGTAAA TAGCACCCAC ACCGGTGGCGG GAATTGGCC 8000
 TATCAATTG CGTAAAAAT AAATAATCAA CAAATGGCA TCGTTTTAA 8050
 TAAAGTGTGATC TATAACCGAAT TCAGCTTTG TTCCCTTTAG TGAGGGTTAA 8100
 15 TTGCGCGCTT GGCATAATCA TGGTCATAGC TGTTCTCTGT TGAAATTGT 8150
 TATCCGCTCA CAATTCCACA CAACATACGA GCCGGAAGCA TAAAGTGTAA 8200
 AGCCTGGGT GCCTAATGAG TGAGCTAATC CACATTAATT GCGTTGCGCT 8250
 CACTGCCGC TTTCAGTCG GGAAACCTGT CGTGCAGCT GCATTAATGA 8300
 ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTCCGC 8350
 20 TTCCCTCGCTC ACTGACTCGC TCGCCTCGGT CGTTCGGCTG CGGCGAGCGG 8400
 TATCAGCTCA CTCAAAGGCG GAAATACGGT TATCCACAGA ATCAGGGAT 8450
 AACGCAGGG AGAACATGTG AGCAAAAGG CAGCAAAGG CCAGGAACCG 8500
 TAAAAAGGCC GCGGTGCTGG CGTTTTCCA TAGGCTCCGC CCCCCCTGACG 8550
 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA 8600
 25 CTATAAAGAT ACCAGGCCTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC 8650
 TGTTCCGACC CTGCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG 8700
 GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG 8750
 TAGGTCGTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC 8800
 CGACCGCTGC GCCTTATCCG GAACTATCG TCTTGAGTCC AACCCGGTAA 8850
 30 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA 8900
 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA 8950
 CGGCTACACT AGAAGGACAG TATTGGTAT CTGCGCTCTG CTGAAGGCCAG 9000
 TTACCTTCGG AAAAGAGATT GGTAGCTCTT GATCCGGCAA ACAAAACCACC 9050
 GCTGGTAGCG GTGGTTTTT TGTTTGCAAG CAGCACATTA CGCGCAGAAA 9100
 35 AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC 9150
 AGTGGAACGA AAACTCACGT TAAGGGATT TGTCATGAG ATTATCAAAA 9200
 AGGATCTTC A CCTAGATCCT TTTAAATTAA AAATGAAGATT TTAAATCAAT 9250
 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA 9300
 GTGAGGCCACC TATCTCAGCG ATCTGTCTAT TTCTTCATC CATAGITGCC 9350
 40 TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG 9400
 CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT 9450
 TATCAGCAAT AAACCAGCCA GCGGAAAGGG CGGAGCGCAG AAGTGGTCTT 9500
 GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG 9550
 AGTAAGTAGT TCGCCAGTTA ATAGTTTGC CGAACGTTGTT GCCATTGCTA 9600
 45 CAGGCATCGT GGTGTACCGC TCGTCGTTG GTATGGCTTC ATTCACTCTCC 9650
 GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGAAAAAA 9700
 AGCGGTTAGC TCCCTCGTC CTCCGATCGT TGTAGAGT AAGTTGGCCG 9750
 CAGTGTATC ACTCATGTT ATGGCAGCAC TGCTAATTCT TCTTACTGTG 9800
 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC 9850
 50 ATTCTGAGAA TAGTGTATGC GGCAGCCAG TTGCTCTTGC CCGCGTCAA 9900
 TACGGSATAA TACCGGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT 9950
 GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG 10000
 ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT 10050
 TTACTTTCAC CAGCGTTCTC GGGTGAGCAA AAACAGGAAG GCAAAATGCC 10100
 GCAAAAAAGG GAATAAGGC GACACGGAAA TGTGAATAC TCATACTCTT 10150
 CCTTTTCAA TATTATTGAA GCATTATCA GGTTTATTGT CTCATGAGCG 10200
 GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC 10250
 ACATTTCCCC GAAAAGTGCC AC 10272

55

SEQ ID NO:31 (pTnMod(Oval/ENT tag/Proins/PA) - Chicken)

60 CTGACCGGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG 50
 CGCACCGTGA CCGCTACACT TGCCAGCGCC CTAGGCCCGC CTCCCTTCGC 100
 TTTCTTCCTC TTCTTCTCG CCACGTTCGC CGGCATCAGA TTGGCTATTG 150
 GCCATTGCTACGTTGATC CATATCATAA TATGTACATT TATATTGGCT 200

	CATGTCCAAC	ATTACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	250
	TAGTAATCAA	TTACGGGTC	ATTAGTCAT	AGCCCATATA	TGGAGTTCCG	300
	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	350
5	CCCGCCCAT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	400
	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT	AAACTGCCA	450
	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTACGCC	CCTATTGACG	500
	TCAATGACGG	TAATGGCCC	GCCTGGCATT	ATGCCAGTA	CATGACCTTA	550
	TGGGACTTTC	CTACTTGCA	GTACATCTAC	GTATTAGTC	TCGCTATTAC	600
	CATGGTATC	CGGTTTGGC	AGTACATCAA	TGGCGTGGA	TAGCGGTTTG	650
10	ACTCACGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG	700
	TTTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGCTGTA	ACAACTCCGC	750
	CCCATTGACG	CAAATGGCG	GTAGCGTGT	ACGGTGGGAG	GTCTATATAA	800
	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACGC	850
15	TGTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGCCG	900
	GGAACGGTGC	ATTGGAACGC	GGATTCCCG	TGCAAGAGT	GACGTAAGTA	950
	CCGCCTATAG	ACTCTATAGG	CACACCCCT	TGGCTCTTAT	GCATGCTATA	1000
	CTGTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTTCCTTAT	GCTATAGGTG	1050
	ATGGTATAGC	TTAGCCTATA	GGTGTGGGTT	ATTGACCTT	ATTGACCACT	1100
20	CCCCTATTGG	TGACGATACT	TTCCATTACT	AATCCATAAC	ATGGCTCTTT	1150
	GCCACAACTA	TCTCTATTGG	CTATATGCCA	ATATCTGTC	CTTCAGAGAC	1200
	TGACACGGAC	TCTGTATTTC	TACAGGATGG	GGTCCCATTT	ATTATTTACA	1250
	AATTACACATA	TACAACAACG	CCGTCACCCG	TGCCCCGAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAATCTCGG	GTACGTGTT	CGGACATGGG	1350
	CTCTTCTCCG	GTAGCGGGCG	AGCTTCCACA	TCCGAGCCCT	GGTCCCATGC	1400
25	CTCCAGCGGC	TCATGGTCGC	TCGGCAGCTC	CTTGCTCTA	ACAGTGGAGG	1450
	CCAGACTTAG	GCACAGCACA	ATGCCACCA	CCACCACTGT	GCCGCACAAG	1500
	GCCGTGGCG	TAGGGTATGT	GTCTGAAAAT	GAGCGTGGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGGAA	GACTTAAGGC	AGCGGCAGAA	GAAGATGCA	1600
	GCAGCTGACT	TGTTGTATTTC	TGATAAGAGT	CAGAGGTAAC	TCCCCTTGCG	1650
30	GTGCTGTTAA	CGGTGGAGGG	CACTGTAGTC	TGAGCAGTAC	TCGTTGCTGC	1700
	CGCGCGGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCCCTT	1750
	CCATGGGTCT	TTTCTGCACT	CACCGTCGGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGATT	CTCTTACCA	ATTCTGCC	GAATTACACT	TAAAACGACT	1850
	CAACAGCTT	ACGTTGGCTT	GCCACGCATT	ACTTGACTGT	AAAACCTCTA	1900
35	CTCTTACCGA	ACTTGGCGT	AACTGCCAA	CCAAACGGAG	AACAAAACAT	1950
	AACATCAAAAC	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAAGA	2000
	GGCAGCTCGT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TCGGGAATAC	2050
	GATGCCATT	GTACTTGTG	ACTGGCTGA	TATTCTGTGAG	AAAAACGAC	2100
	TTATGGTATT	GCGAGCTCA	GTGCCACTAC	ACGGTCGTT	TGTTACTCTT	2150
40	TATGAGAAAG	CGTTCCCGCT	TTCAGAGCA	TGTTCAAAGA	AAGCTCATGA	2200
	CCAATTCTA	GCCGACCTTG	CGAGCATTCT	ACCGAGTAAC	ACCACACCGC	2250
	TCATTGTCAG	TGATGCTGC	TTTAAAGTC	CATGGTATAA	ATCCGTTGAG	2300
	AAGCTGGTT	GGTACTGGTT	AACTCGAGTA	AGAGGAAAG	TACAATATGC	2350
	AGACCTAGGA	GCGGAAACACT	GGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
45	CATCTAGTC	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAAGCAAT	2450
	CCAATCTCAT	GCCAAATTCT	ATTGTATAAA	TCTCGCTTA	AAGGCCGAAA	2500
	AAATCAGCGC	TCGACACCGA	CTCATTGTCA	CCACCCGTCA	CCTAAAATCT	2550
	ACTCAGCGTC	GGCAAAGGAG	CCATGGGTT	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTGAA	CACCCAAACA	ACTTGTAAAT	ATCTATTGCA	AGCGAATGCA	2650
50	GATTGAAGAA	ACCTTCCGAG	ACTTGAAAG	TCCTGCCAAC	GGACTAGGCC	2700
	TACGCCATAG	CGAACGAGC	AGCTCAGAGC	GTTCATGATAT	CATGCTGCTA	2750
	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	CTTGCAGGGCG	TTCATGCTCA	2800
	GAAACAAGGT	TGGGACAAGC	ACTTCACCGC	TAACACAGTC	AGAAATCGAA	2850
	ACGTACTCTC	AACAGTTCGC	TTAGGCATGG	AAGTTTGC	GCATTCTGGC	2900
55	TACACAATAA	CAAGGGAAAGA	CTTACTCGTG	GCTGCAACCC	TACTAGCTCA	2950
	AAATTTATTTC	ACACATGGTT	ACGCTTGGG	GAAATTATGA	TAATGATCCA	3000
	GATCACTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTGTGTTGG	3050
	TTTTTGTGG	ATCTGCTGT	CCTTCTAGTT	GCCAGCCATC	TGTTGTTTGC	3100
	CCCTCCCCCG	TGCCTTCCTT	GACCTGGAA	GGTGCACACTC	CCACTGTCCT	3150
60	TTCCTAAATAA	AATGAGGAAA	TTGCATCGCA	TTGTCTGAGT	AGGTGTCATT	3200
	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACA	GCAAGGGGAA	GGATTGGGAA	3250

	GACAATAGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAAGG	TGCTGAAGAA	3400
5	TTGACCCGGT	GACCAAAGGT	GCCTTTATC	ATCACTTTAA	AAATAAAAAAA	3450
	CAATTACTCA	GTGCCTGTTA	TAAGCAGCAA	TTAATTATGA	TTGATGCCTA	3500
	CATCACAACA	AAAAGTGATT	TAACAAATGG	TTGGTCTGCC	TTAGAAAGTA	3550
	TATTTGAACA	TTATCTTGAT	TATATTATTG	ATAATAATAA	AAACCTTATC	3600
	CCTATCCAAG	AAAGTGATGCC	TATCATGGT	TGGAATGAC	TTGAAAAAAA	3650
10	TTAGCCTTGA	ATACATTACT	GGTAAGGTAA	ACGCCATTGT	CAGCAAATG	3700
	ATCCAAGAGA	ACCAACTAA	AGCTTCTTG	ACGGAATGTT	AATTCTCGTT	3750
	GACCCCTGAGC	ACTGATGAAT	CCCCATATGA	TTTTGGTAAA	AATCATTAAAG	3800
	TTAAGGTGGA	TACACATCTT	GTCTATGAT	CCCGGTAATG	TGAGTTAGCT	3850
	CACTCATTAG	GCACCCCCAGG	CTTTACACTT	TATGCTTCCG	GCTCGTATGT	3900
15	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTC	ACACAGGAAA	CAGCTATGAC	3950
	CATGATTAGC	CCAAGCGCGC	AATTAACCC	CACTAAAGGG	ACAAAAAGCT	4000
	GGAGGCTCCAC	CGCGGTGGCG	GGCGCTCTAG	AACTAGTGG	TCCCCCGGGG	4050
	AGGTCAAGAT	GGTTTCTTA	CTGTTGTCA	ATTCTTAT	TTCAATACAG	4100
	AACAATAGCT	TCTATAACTG	AAATATATT	GCTATTGTAT	ATTATGATTG	4150
20	TCCCTCGAAC	CATGAACACT	CCTCCAGCTG	AATTTCACAA	TTCCCTGTGTC	4200
	ATCTGCCAGG	CCATTAAAGTT	ATTCATGAA	GATCTTTGAG	GAACACTGCA	4250
	AGTTCATATC	ATAAAACACAT	TTGAAATTGA	GTATTGTTT	GCATTGTATG	4300
	GAGCTATGTT	TTGCTGTATC	CTCAGAAAAA	AAAGTTGTTA	TAAAGCATTC	4350
	ACACCCATAA	AAAGATAGAT	TTAAATATT	CAGCTATAGG	AAAGAAAGTG	4400
25	CGTCTGCTCT	TCACTCTAGT	CTCAGTTGGC	TCCTTCACAT	GCATGCTTCT	4450
	TTATTTCTCC	TATTTGTCA	AGAAAATAAT	AGGTCACCTC	TTGTTCTCAC	4500
	TTATGCTCTG	CCTAGCATGG	CTCAGATGCA	CGTTGTAGAT	ACAAGAAGGA	4550
	TCAAATGAAA	CAGACTCTG	GTCTGTTACT	ACAACCATAG	TAATAAGCAC	4600
	ACTAACTAAT	AATTGCTAAT	TATGTTTCC	ATCTCTAAGG	TTCCACATT	4650
30	TTTCTGTTT	CTTAAAGATC	CCATTATCTG	GTGTTAACTG	AAGCTCAATG	4700
	GAACATGAGC	AATATTTCC	AGTCTCTCT	CCCATCCAAC	AGTCCTGATG	4750
	GATTAGCAGA	ACAGGCAGAA	AACACATTGT	TACCCAGAAT	AAAAAACTAA	4800
	TATTTGCTCT	CCATTCAATC	CAAATGGAC	CTATTGAAAC	AAAAATCTAA	4850
	CCCAATCCCA	TTAAATGATT	TCTATGGCGT	CAAAGGTCAA	ACTTCTGAAG	4900
35	GGAACCTGTG	GGTGGGTAC	AATTCAAGGT	ATATATCCC	CAGGGCTCAG	4950
	CGGATCCATG	GGCTCCATCG	GGCGCAGCAAG	CATGGAATT	TGTTTTGATG	5000
	TATTCAAGGA	GCTCAAAGTC	CACCATGCC	ATGAGAACAT	CTTCTACTGC	5050
	CCCATTGCCA	TCATGTCAGC	TCTAGCCATG	GTATACCTGG	GTGCAAAGAGA	5100
	CAGCACCAGG	ACACAGATAA	ATAAGTTGT	TCGCTTTGAT	AAACTTCCAG	5150
40	GATTGGAGA	CACTATTGAA	GCTCAGTGTG	GCACATCTGT	AAACGTTCAC	5200
	TCTTCACCTA	GAGACATCT	CAACCAAATC	ACCAAACCAA	ATGATGTTA	5250
	TCGTTTCAGC	CTTGCAGTA	GAECTTATGC	TGAAGAGAGA	TACCCAATCC	5300
	TGCCAGAATA	CTTGCAGTGT	GTGAAGAAC	TGTATAGAGG	AGGCTTGGAA	5350
	CCTATCAACT	TTCAAAACAGC	TGCAAGATCAA	GGCAGAGAGC	TCATCAATTC	5400
45	CTGGGTAGAA	AGTCAGACAA	ATGGAATTAT	CAGAAATGTC	CTTCAGCCAA	5450
	GCTCCGTGGA	TTCTCAAAC	GCAATGGTT	TGTTTAATGC	CATTGTCTTC	5500
	AAAGGACTGT	GGGAGAAAAC	ATTTAAGGAT	GAAGACACAC	AAGCAATGCC	5550
	TTTCAGAGTG	ACTGAGCAAG	AAAGCAACCC	TGTGCAGATG	ATGTACCAAGA	5600
	TTGGTTTATT	TAGAGTGGCA	TCAATGGCTT	CTGAGAAAAT	GAAGATCCTG	5650
50	GAGCTTCCAT	TTGCCAGTGG	GACAATGAGC	ATGTTGGTGC	TGTTGCCTGA	5700
	TGAAGTCTCA	GGCCTTGAGC	AGCTTGAGAG	TATAATCAAC	TTTGAACAC	5750
	TGACTGAATG	GACCAAGTCT	AATGTTATGG	AAAGAGAGGAA	GATCAAAGTG	5800
	TACTTACCTC	GCATGAAGAT	GGAGGAAAAA	TACAACCTCA	CATCTGTCTT	5850
	AATGGCTATG	GGCATTACTG	ACGTTTTAG	CTCTTCAGCC	AATCTGTCTG	5900
55	GCATCTCCTC	AGCAGAGAGC	CTGAAGATAT	CTCAAGCTGT	CCATGCAGCA	5950
	CATGCAGAAA	TCAATGAAGC	AGGCAGAGAG	GTGGTAGGGT	CAGCAGAGGC	6000
	TGGAGTGGAT	GCTGCAAGCG	TCTCTGAAGA	ATTAGGGCT	GACCATCCAT	6050
	TCCTCTCTG	TATCAAGCAC	ATCGCAACCA	ACGCCGTCT	CTTCTTTGGC	6100
	AGATGTGTTT	CCCCCTCGCG	GCCAGCAGAT	GACGCACCAAG	CAGATGACGC	6150
60	ACCAGCAGAT	GACGCACCAAG	CAGATGACGC	ACCAAGCAGAT	GACGCACCAAG	6200
	CAGATGACGC	AACAACATGT	ATCCGTAAAG	GCTCTTGTTG	CTGGATCGGC	6250
	CTGCTGGATG	ACGATGACAA	ATTGTGAAC	CAACACCTGT	GGGGCTCACCA	6300

CCTGGTGGAA GCTCTCTACC TAGTGTGCGG GGAACGAGGC TTCTTCTACA 6350
 CACCCAAGAC CCGCCGGGAG GCAGAGGACC TGCAAGTGGG GCAGGTGGAG 6400
 CTGGCGGGG GCCCCTGGC AGGCAGCCTG CAGCCCTTGG CCCCTGGAGGG 6450
 5 GTCCCTGCAG AAGCGTGGCA TTGTGGAACA ATGCTGTACC AGCATCTGCT 6500
 CCCTCTACCA GCTGGAGAAC TACTGCAACT AGGGCCCTG GATCCAGATC 6550
 ACTTCTGGCT AATAAAAAGAT CAGAGCTCTA GAGATCTGTG TGTTGGTTT 6600
 TTGTGGATCT GCTGTGCCCT CTAGTGTCCA GCCATCTGTT GTTGTGCCCT 6650
 CCCCGTGCC TTCCCTGACC CTGGAAGGTG CCACTCCAC TGTCCCTTCC 6700
 TAATAAAATG AGGAAATTGC ATCGCATTGTG CTGAGTAGGT GTCATTCTAT 6750
 10 TCTGGGGGT GGGGTGGGGC AGCACAGCAA GGGGGAGGAT TGGGAAGACA 6800
 ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT CTATGGGTAC CTCTCTCT 6850
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCTCT CTCGAGGGGG 6900
 GGCCCGGTAC CCAATTGCCC CTATAGTGG AG TCGTATTACG CGCGCTCACT 6950
 GGCGTGTGTT TTACAACGTC GTGACTGGGA AAACCCCTGGC GTTACCCAAC 7000
 15 TTAATCGCCT TGCAGCACAT CCCCCCTTCG CCAGCTGGCG TAATAGCGAA 7050
 GAGGGCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA 7100
 ATGGAAATTG TAAGCGTTAA TATTTGTTA AAATTCGCGT TAAATTTTG 7150
 TTAAATCAGC TCATTTTTA CCAATAGGC CGAACATCGGC AAAATCCCTT 7200
 ATAAATCAA AGAATAGACC GAGATAGGGT TGAGTGTGTT TCCAGTTGG 7250
 20 AACAAAGAGTC CACTATTTAA GAACGTGGAC TCCAACGTC AAGGGCGAAA 7300
 AACCGTCTAT CAGGGCGATG GCCCACTACT CCGGGATCAT ATGACAAGAT 7350
 GTGTATCCAC CTTAACTTAA TGATTTTAC CAAAATCATT AGGGGATTCA 7400
 TCAGTGCTCA GGGTCAACGA GAATTAACAT TCCGTCAAGGA AAGCTTATGA 7450
 TGATGATGTG CTTAAAAGT TACTCAATGG CTGGTTATGC ATATCGCAAT 7500
 25 ACATGCGAAA AACCTAAAAG AGCTTGCCGA TAAAAAAGGC CAATTTATIG 7550
 CTATTTACCG CGGCTTTTA TTGAGCTTGA AAGATAAAATA AAATAGATAG 7600
 GTTTTATTG AAGCTAAATC TTCTTTATCG TAAAAAATGC CCTCTTGGGT 7650
 TATCAAGAGG GTCATTATAT TTCCGGAAT AACATCATT GGTGACGAAA 7700
 TAACTAAGCA CTTGTCTCT GTTTACTCCC CTGAGCTTGA GGGGTTAAC 7750
 30 TGAAGGTCA CGATAGCAGG ATAATAATAC AGTAAAACGC TAAACCAATA 7800
 ATCCAAATCC AGCCATCCCA AATTGGTAGT GAATGATTAT AAATAACAGC 7850
 AAACAGTAAT GGGCCAATAA CACCGGTTGC ATTGGTAAGG CTCACCAATA 7900
 ATCCCTGTAA AGCACCTTGC TGATGACTCT TTGTTGGAT AGACATCACT 7950
 CCTGTAAATG CAGGTAAGC GATCCACCA CCAGCCAATA AAATTAAAAC 8000
 35 AGGGAAAAGT AACCAACCTT CAGATATAA CGCTAAAGG GCAAATGCAC 8050
 TACTATCTGC AATAAAATCCG AGCACTACTG CCGTTTTTC GCCCCATTAA 8100
 GTGGCTATTG TTCCCTGCCAC AAAGGCTTGG AATACTGAGT GTAAAAGACC 8150
 AAGACCCGCT AATGAAAAGC CAACCATCAT GCTATTCCAT CCAAAACGAT 8200
 TTTCGTAAA TAGCACCCAC ACCGTTGCGG GAATTGGCC TATCAATTGC 8250
 40 GCTGAAAAT AAATAATCAA CAAATGGCA TCGTTTTAA TAAAGTGATG 8300
 TATACCGAAT TCAGCTTTG TTCCCTTTAG TGAGGGTTAA TTGCGCGCTT 8350
 GCGCTAATCA TGGTCATAGC TGTTCCCTGT GTGAAATTGT TATCCGCTCA 8400
 CAATTCCACACAAACATACGA GCGGAAAGCA TAAAGTGTAA AGCCTGGGGT 8450
 GCCTTAATGAG TGAGCTAATC CACATTAATT GCGTTGCCT CACTGCCCGC 8500
 45 TTTCCAGTCG GGAAACCTGT CGTGCAGCT GCATTAATGA ATCGGCCAAC 8550
 GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTCCGC TTCCCTCGCTC 8600
 ACTGACTCGC TCGCCTCGGT CGTTGGCTG CGGGAGCGG TATCAGCTCA 8650
 CTCAAAGCG GTAAATACGGT TATCCACAGA ATCAGGGAT AACGCAGGAA 8700
 AGAACATGTG AGCAAAAGGC CAGCAAAGG CCAGGAACCG TAAAAGGCC 8750
 50 GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCCTGACG AGCATCACAA 8800
 AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAAGAT 8850
 ACCAGGGCTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCGGACC 8900
 CTGCGCTTA CGGATACCT GTCCGCTTT CTCCCTTCGG GAAGCGTGGC 8950
 GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTGGTG TAGGTGTTTC 9000
 55 GCTCCAAGCT GGGCTGTGTC CACGAACCCC CCAGTCAGCC CGACCGCTGC 9050
 GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT 9100
 ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG 9150
 TAGGCGGTGC TACAGAGTT TGTAAAGTGGT GGCCTAACTA CGGCTACACT 9200
 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGGCCAG TTACCTTCGG 9250
 60 AAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAAACCACC GCTGGTAGCG 9300
 GTGGTTTTT TGTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT 9350

	CAAGAAGATC	CTTGATCTT	TTCTACGGGG	TCTGACGCTC	AGTGGAAACGA	9400
	AAACTCACGT	TAAGGGATTT	TGGTCATGAG	ATTATCAAAA	AGGATCTTC	9450
	CCTAGATCCT	TTTAAATTAA	AAATGAAGTT	TTAAATCAAT	CTAAAGTATA	9500
5	TATGAGTAA	CTTGGTCTGA	CAGTTACCAA	TGCTTAATCA	GTGAGGCACC	9550
	TATCTCAGCG	ATCTGTCTAT	TTCGTTCATC	CATAGTTGCC	TGACTCCCCG	9600
	TCGTGTAGAT	AACTACGATA	CGGGAGGCT	TACCATCTGG	CCCCAGTGCT	9650
	GCAATGATAC	CCCGAGACCC	ACCGTCACCG	GCTCCAGATT	TATCAGCAAT	9700
	AAACCAGCCA	CCCGGAAGGG	CCGACGCGAG	AAGTGGTCT	GCAACTTTAT	9750
10	CCGCCTCCAT	CCAGTCTATT	AATTGGTGC	GGGAAGCTAG	AGTAAGTAGT	9800
	TCGCCAGTTA	ATAGTTGCG	CAACGTTGTT	GCCATTGCTA	CAGGCATCGT	9850
	GGTGTACCGC	TCGTCGTTG	GTATGGCTTC	ATTCAAGCTCC	GGTTCCCAAC	9900
	GATCAAGGGC	AGTTACATGA	TCCCCATGT	TGTGCAAAAA	AGCGGTTAGC	9950
	TCCTTCGGTC	CTCCGATCGT	TGTCAAGT	AAGTTGGCCG	CAGTGTATC	10000
15	ACTCATGGTT	ATGGCAGCAC	TGCATAATT	TCTTAUTGTC	ATGCCATCCG	10050
	TAAGATGCTT	TTCTGTGACT	GGTGAAGTACT	CAACCAAGTC	ATTCTGAGAA	10100
	TAGTGTATGC	GGCGACCGAG	TTGCTCTTG	CCGGCGTCAA	TACGGGATAAA	10150
	TACCCCGCCA	CATAGCAGAA	CTTTAAAAGT	GCTCATCATT	GGAAAACGTT	10200
	CTTCGGGGCG	AAAATCTCA	AGGATCTTAC	CGCTGTTGAG	ATCCAGTTCG	10250
20	ATGTAACCCA	CTCGTGCACC	CAACTGATCT	TCAGCATCTT	TTACTTTCAC	10300
	CAGCGTTCT	GGGTGAGCAA	AAACAGGAAG	GCAAAATGCC	GCAAAAAAGG	10350
	GAATAAGGGC	GACACGGAAA	TGTTGAAATAC	TCATACTCTT	CCTTTTCAA	10400
	TATTAATTGAA	GCATTTATCA	GGGTTATTGT	CTCATGAGCG	GATACATATT	10450
	TGAATGTATT	TAGAAAAATA	AACAAATAGG	GGTCCCGC	ACATTCCCC	10500
25	AAAAAGTGC	AC				10512

SEQ ID NO: 32 (pTnMod(Oval/ENT tag/Proins/PA) - QUAIL)

30	CTGACGCGCC	CTGTAGCGGC	GCATTAAGCG	CGGCGGGGT	GGTGGTTACG	50
	CGCAGCGTGA	CCGCTACACT	TGCCAGCGCC	CTAGCGCCCG	CTCCTTTCGC	100
	TTCTTCCTC	TCCTTCTCG	CCACGTTCG	CGGCATCAGA	TTGGCTATTG	150
	GCCATTGCAT	ACGTTGTATC	CATATCATAA	TATGTACATT	TATATTGGCT	200
	CATGTCCAAC	ATTACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	250
	TAGTAATCAA	TTACGGGTC	ATTAGTTCAT	AGCCCCATATA	TGGAGTTCCG	300
35	CGTTACATCAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	350
	CCCGCCCCATT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	400
	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT	AAACTGCCA	450
	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAAGTACGCC	CCTATTGACG	500
40	TCAATGACGG	TAATGGCCC	GCCTGGCATT	ATGCCAGTA	CATGACCTTA	550
	TGGGACTTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC	600
	CATGGTGATG	CGGTTTGGC	AGTACATCAA	TGGGCGTGG	TAGCGGTTTG	650
	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG	700
	TTTTCGCACC	AAAATCAACG	GGACTTTCCA	AAATGCTGA	ACAACCTCCG	750
45	CCCATTTGACG	CAAATGGCG	GTAGCGTGT	ACGGTGGAG	GTCTATATAA	800
	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACCG	850
	TGTTTGACG	TCCATAGAAG	ACACCGGAC	CGATCCAGCC	TCCGCGGCCG	900
	GGAACCGGTG	ATTGGAACCG	GGATTCCCCG	TGCCAAGAGT	GACGTAAGTA	950
50	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA	1000
	CTGTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTTCCTTAT	GCTATAGGTG	1050
	ATGGTATAGC	TTAGCCTATA	GGTGTGGTT	ATTGACCATT	ATTGACCACT	1100
	CCCCTATTGG	TGACGATACT	TTCCATTACT	AAATCCATAAC	ATGGCTCTTT	1150
	GCCACAACTA	TCTCTATTGG	CTATATGCCA	ATACTCTGTC	CTTCAGAGAC	1200
	TGACACGGAC	TCTGTATTT	TACAGGATGG	GGTCCCATTT	ATTATTTACA	1250
55	AATTACATA	TACAACAAACG	CCGTCACCG	TGCCCGCAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAATCTGG	GTACGTGTC	CGGACATGGG	1350
	CTCTTCTCCG	GTAGCGGCCG	AGCTTCCACA	TCCGAGCCCT	GGTCCCAGTC	1400
	CTCCAGCGGC	TCATGGTCGC	TCGGCAGCTC	CTTGCTCCTA	ACAGTGGAGG	1450
	CCAGACTTAG	GCACAGCACA	ATGCCACCA	CCACCACTGT	GCCGCACAAG	1500
60	GCCGTGGCG	TAGGGTATGT	GTCTAAAAT	GAGCGTGGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGGAA	GACTTAAGGC	AGCGGCAGAA	GAAGATGCAG	1600
	GCAGCTGAGT	TGTTGTATTC	TGATAAGAGT	CAGAGGTAAC	TCCCAGTTGCG	1650

GTGCTGTTAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TCGTTGCTGC 1700
 CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCCCTT 1750
 CCATGGGTCT TTTCTGCACT CACCGTCGGA CCATGTGTGA ACTTGATATT 1800
 TTACATGATT CTCTTACCA ATTCTGCCCC GAATTACACT TAAAACGACT 1850
 5 CAACAGCTTA ACGTGGCTT GCCACGCATT ACTTGACTGT AAAACTCTCA 1900
 CTCTTACCGA ACTTGGCCGT AACCTGCCA CCAAAGCGAG AACAAAACAT 1950
 AACATCAAAC GAATCGACCG ATTGTTAGGT AATCGTCACC TCCACAAAGA 2000
 GCGACTCGCT GTATACCGTT GGCGATGCTAG CTTTATCTGT TCGGGAATAC 2050
 GATGCCCCATT GTACTGTTG ACTGGTCTGA TATTCTGTGAG CAAAACGAC 2100
 10 TTATGGTATT GCGAGCTTCA GTCGCACTAC ACCTGCGTTC TGTTACTCTT 2150
 TATGAGAAAG CGTCCCCTGCT TTCAAGAGCAA TGTTCAAAGA AAGCTCATGA 2200
 CCAATTCTCA GCCGACCTTG CGAGCATTCT ACCGAGTAAC ACCACACCGC 2250
 TCATTGTCAG TGATGCTGGC TTTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
 AAGCTGGTTT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
 15 AGACCTAGGA GCGGAAAAGT GGAAACCTAT CAGCAACTTA CATGATATGT 2400
 CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAAGCAAT 2450
 CCAATCTCAT GCCAAATCTC ATTGTATAAA TCTCGCTCTA AAGGCCGAAA 2500
 AAATCAGCGC TCGACACCGA CTCAATGTCA CCACCCGTCA CCTAAAATCT 2550
 ACTCAGCGTC GGCAAGGGAG CCATGGTTC TAGCAACTAA CTTACCTGTT 2600
 20 GAAATTGAA CACCCAAACA ACTTGTAAAT ATCTATTGCA AGCGAATGCA 2650
 GATTGAAGAA ACCTTCCGAG ACTTGAAAAG TCCTGCTTAC GGACTAGGCC 2700
 TACGCCATAG CCGAACGAGC AGCTCAGAGC GTTTGATAT CATGCTGCTA 2750
 ATCGCCCTGA TGCTTCAACT AACATGTTGG CTGCGGGCG TTCATGCTCA 2800
 GAAACAAGGT TGGGACAAGC ACTTCCAGGC TAACACAGTC AGAAATCGAA 2850
 25 ACGTACTCTC AACAGTTCGC TTAGGCATGG AAGTTTGCG GCATTCTGGC 2900
 TACACAATAA CAAGGGAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTATTC ACACATGGTT ACGCTTTGGG GAAATTATGA TAATGATCCA 3000
 GATCACTCTC GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 TTTTGTGGA ATCTGCTGTG CCTCTCTAGTT GCCAGGCCATC TGTTGTTTC 3100
 30 CCCTCCCCCG TGCCCTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT 3150
 TTCCATAATAA AATGAGGAAA TTGCAATCGCA TTGCTCTGAGT AGGTGTCATT 3200
 CTATTCTGGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGG GGATTGGGAA 3250
 GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CCGTACAGG TGCTGAAGAA 3400
 35 TTGACCCGGT GACCAAAGGT GCCTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCCCTGTTA TAAGCAGCAA TTAATTATGA TTGATGCCCTA 3500
 CATCACAAACA AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTGAAACA TTATCTTGAT TATATTATTG ATAATAATAA AAACCTTATC 3600
 40 CCTATCCAAG AAGTGATGCC TATCATTGGT TGAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAGGTAAC ACGCCATTGT CAGCAAATG 3700
 ATCCAAGAGA ACCAACTAA AGCTTCTTG ACAGGAATGTT AATTCTCGTT 3750
 GACCCCTGAGC ACTGATGAAT CCCCTATGA TTGGTAAAGG AATCATTAAAG 3800
 TTAAGGTGGA TACACATCTT GTCATATGAT CCCGGTAATG TGAGTTAGCT 3850
 45 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGACGGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGCGC AATTAACCCCT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTAGTGGG TCCCCGGGG 4050
 AGGTCAAGAT GGGTTCTTA CTGTTGTCA ATTCTATTAT TTCAATACAG 4100
 50 AACAAAAGCT TCTATAACTG AAATATATTG GCTATTGTAT ATTATGATTC 4150
 TCCCTCGAAC CATGAAACT CCTCCAGCTG AATTTCACAA TTCTCTGTC 4200
 ATCTGCCAGG CTGGAAGATC ATGGAAGATC TCTGAGGAAC ATTGCAAGTT 4250
 CATACCATAA ACTCATTGTT AATTGAGTAT TATTTGCTT TGAATGGAGC 4300
 TATGTTTTGC AGTTCCCTCA AAAGAAAAGC TTGTTATAAA GCGTCTACAC 4350
 55 CCATCAAAG ATATATTAA ATATTCAAC TACAGAAAAGA TTGTTCTGC 4400
 TCTTCACTCT GATCTCAGTT GGTTCTTCA CGTACATGCT TCTTTATTG 4450
 CCTATTTGT CAAGAAAATA ATAGGTCAAG TCCTGTTCTC ACTTATCTCC 4500
 TGCCTAGCAT GGCTTAGATG CACGGTGTAC ATTCAAGAAG GATCAAATGA 4550
 AACAGACTTC TGGTCTGTTA CAACAACCAT AGTAATAAAC AGACTAACTA 4600
 60 ATAATTGCTA ATTATGTTT CCATCTCTAA GGTTCCCACCA TTTTTCTGTT 4650
 TTAAGATCCC ATTATCTGGT TGTAAGTGAAC GCTCAATGGA ACATGAACAG 4700

TATTTCTAG TCTTTCTCC AGCAATCCTG ACGGATTAGA AGAACTGGCA 4750
 GAAAACACTT TGTTACCCAG AATTAACAAAC TAATATTGC TCTCCCTTC 4800
 ATCCAAAATG GACCTATTGA AACTAAAATC TGACCCAATC CCATTAAATT 4850
 ATTTCTATGG CGTCAAAGGT CAAACTTTG AAGGGAACTT GTGGGTGGGT 4900
 5 CCCAATTCTAG GCTATATATT CCCACAGGGCT CAGCCAGTGG ATCCATGGC 4950
 TCCATCGGTG CAGCAAGCAT GGAATTGTG TTTGATGTAT TCAAGGAGCT 5000
 CAAAGTCCAC CATGCCAATG ACAACATGCT CTACTCCCCC TTTGCCATCT 5050
 TGTCAACTCT GCCCATGGTC TTCTAGGTG CAAAGACAG CACCAGGACC 5100
 CAGATAATAA AGGTTGTTCA CTTTGATAAA CTTCAGGAT TCGGAGACAG 5150
 10 TATTGAGCT CAGTGTGGCA CATCTGAA TGTGACTCT TCACTTAGAG 5200
 ACATACTCAA CCAAATCACC AAACAAAATG ATGCTTATTG GTTCAGCCTT 5250
 GCCAGTAGAC TTTATGCTCA AGAGACATAC ACAGCTGTG CGGAATACTT 5300
 GCAATGTGTG AAGGAACGT ATAGAGGAGG CTTAGAATCC GTCAACTTTC 5350
 15 AAACAGCTGC AGATCAAGCC AGAGGCTCA TCAATGCCTG GGTAGAAAGT 5400
 CAGACAAACG GAATTATCTAG AACATCCTT CAGCCAACT CGGTGGATT 5450
 TCAAACATGCA ATGGCTCTGG TTAATGCCAT TGCCTTCAAG GGACTGTGGG 5500
 AGAAAGCATT TAAGGCTGAA GACACGCAA CAATACCTTT CAGAGTGACT 5550
 GAGCAAGAAA GCAAACCTGT CGAGATGATG TACCAAGATG GTTCATTAA 5600
 AGTGGCATCA ATGGCTCTGG AGAAAATGAA GATCCTGGAG CTTCCATTG 5650
 20 CCAGTGGAAC AATGAGCATG TTGGTGTGTG TGCCCTGATGA TGTCTCAGGC 5700
 CTTGAGCAGC TTGAGAGTAT AATCAGCTT GAAAAACTGA CTGAATGGAC 5750
 CAGTTCTAGT ATTATGGAAG AGAGGAAGGT CAAAGTGTAC TTACCTCGCA 5800
 TGAAGATGGA GGAGAAATAC AACCTCACAT CTCTCTTAAT GGCTATGGGA 5850
 ATTACTGACC TGTCAGCTC TTCAAGCAAT CTGTCTGGCA TCTCCTCAGT 5900
 25 AGGGAGCCTG AAGATATCTC AAGCTGTCCA TGAGCACAT GCAGAAATCA 5950
 ATGAAGCGGG CAGAGATGTG TAGGGCTCAG CAGAGGCTGG AGTGGATGCT 6000
 ACTGAAGAAT TTAGGGCTGA CCATCCATTC CTCTCTGTG TCAAGCACAT 6050
 CGAAACCAAC GCCATTCTCC TCTTGGCAG ATGTTGTTCT CCGCGGCCAG 6100
 CAGATGACGC ACCAGCAGAT GACCCACCAAG CAGATGACGC ACCAGCAGAT 6150
 30 GACGCACCAAG CAGATGACGC ACCAGCAGAT GACGCAACAA CATGTATCCT 6200
 GAAAGGCTCT TGTGGCTGGA TCGGCTGCT GGATGACGAT GACAAATTG 6250
 TGAACCAACA CCTGTGGCAGC TCACACCTGG TGAAGCTCT CTACCTAGTG 6300
 TGCGGGGAAAC GAGGCTCTT CTACACACCC AAGACCCGCC GGGAGGCAGA 6350
 GGACCTGCAAG GTGGGGCAGG TGGAGCTGGG CGGGGGCCCT GGTGCAGGCA 6400
 35 GCCTGCAGCC CTTGGCCCTG GAGGGTCCC TGAGAACCCG TGGCATTGTG 6450
 GAACAATGCT GTACAGCAT CTGCTCCCTC TACCAAGCTGG AGAAACTACTG 6500
 CAACTAGGGC GCCTGGATCC AGATCACCTC TGCTTAATAA AAGATCAGAG 6550
 CTCTAGAGAT CTGTGTGTTG GTTTTTGATGCTGTG TGACCTCTAGT 6600
 TGCCAGCCAT CTGTTGTTG CCCCTCCCCC TGCCCTTCCCT TGACCCCTGGA 6650
 40 AGGTGCCACT CCCACTGTCC TTCTCTAATA AAATGAGGAA ATTGCATCGC 6700
 ATTGTCTGAG TAGGTGTATC TCTATTCTGG GGGGTGGGGT GGGGCAGCAC 6750
 AGCAAGGGGG AGGATTGGA AGACAATAGC AGGCATGTG GGGATGCGGT 6800
 GGGCTCTATG GTTACCTCTC TCTCTCTC TCTCTCTC TCTCTCTC 6850
 TCTCTCGGTG CCTCTCTCGA GGGGGGCCC GTTACCAAT TCGCCCTATA 6900
 45 GTGAGTCGTA TTACGCGCGC TCACTGGCCG TCGTTTACA ACGTCGTGAC 6950
 TGGGAAAACC CTGGCGTTAC CCAACTTAAT CGCCCTTGAG CACATCCCCC 7000
 TTTCGCCAGC TGGCGTAATA GCGAAGAGGC CGGCACCGAT CGCCCTTCCC 7050
 AACAGTTGCC CAGCCTGAAT GGCAGATGGA AATTGTAAGC GTTAATATT 7100
 TGTTAAATT CGCGTTAAAT TTTGTTAAA TCAGCTCATT TTTAACCAA 7150
 50 TAGGCCGAAAT CGGCCAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT 7200
 AGGGTTGAGT GTTGTTCAG TTTGGAACAA GAGTCCACTA TTAAAGAACG 7250
 TGGACTCCAA CGTCAAAGGG CGAAAACCG TCTATCAGGG CGATGGCCCA 7300
 CTACTCCGGG ATCATATGAC AAGATGTA TCCACCTTAA CTTAATGATT 7350
 TTTACCAAAAT TCATTAGGGG ATTCACTAGT GCTCAGGGTC AACGAGAATT 7400
 55 AACATTCCGT CAGGAAAGCT ATGATGATG ATGTCCTTAA AAACTTACTC 7450
 AATGGCTGGT TATGCATATC GCAATACATG CGAAAACCTT AAAAGAGCTT 7500
 GCCGATAAAA AAGGCCAATT TATTGCTATT TACCGCGGCT TTTTATTGAG 7550
 CTTGAAAGAT AAATAAAATA GATAGGTTT ATTGGAAGCT AAATCTTCTT 7600
 TATCGAAAAA AATGCCCTCT TGGTTATCA AGAGGGTCAT TATATTTCGC 7650
 60 GGAATAACAT CATTGGTGA CGAATAACT AAGCACTTGT CTCCCTGTTA 7700
 CTCCCCCTGAG CTTGAGGGGT TAACATGAAG GTCACTCGATA GCAGGATAAT 7750

AATAACAGTAA AACGCTAAAC CAATAATCCA AATCCAGCCA TCCCAAATTG 7800
 GTAGTGAATG ATTATAAATA ACAGCAAACA GTAATGGGCC AATAACACCG 7850
 GTTGCATTGG TAAGGCTCAC CAATAATCCC TGAAAGCAC CTTGCTGATG 7900
 5 ACTCITTGTT TGGATAGACA TCACTCCCTG TAATGCAAGGT AAAGCGATCC 7950
 CACCAACCAGC CAATAAAATT AAAACAGGGA AACTAACCA ACCTTCAGAT 8000
 ATAAACGCTA AAAAGGCAAA TGCACTACTA TCTGCAATAA ATCCGAGCAG 8050
 TACTGCCGTT TTTTCGCCCTT ATTAGTGGC TATTCCTCTC GCCACAAAGG 8100
 CTTGGAATAC TGAGTGTAAA AGACCAAGAC CCGCTAATGA AAAGCCAACC 8150
 10 ATCATGCTAT TCCATCCAAA ACGATTTCG GTAAATAGCA CCCACACCGT 8200
 TGGGGAAATT TGGCCTATCA ATTGGCTGA AAAATAAATA ATCAACAAA 8250
 TGGCATCGTT TAAATAAAG TGATGTATAC CGAATTACGC TTTTGTCCC 8300
 TTTAGTGAGG GTTAATTGCG CGCTTGGCGT AATCATGGTC ATAGCTGTT 8350
 CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG 8400
 15 AAGCATAAAAG TGTAAAGCCT GGGGTGCTA ATGAGTGAGC TAACTCACAT 8450
 TAATTGCGTT GCGCTCACTG CCCGCTTCC AGTCGGAAA CCTGTCGTG 8500
 CAGCTGCATT AATGAATCGG CCAACCGCG GGGAGAGGCG GTTTGCGTAT 8550
 TGGGCGCTCT TCCGCTCTC CGCTCACTGA CTGCGTGC CG 8600
 GGCTCCGGCG AGCGGTATCA GCTCACTCAA AGCGGGTAAT ACGGTTATCC 8650
 ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA 8700
 20 AAAGGCCAGG AACCGTAAAA AGGCCCGTT GCTGGCGTTT TTCCATAGGC 8750
 TCCGCCCCCC TGACGAGCAT CACAAAATC GACGCTCAAG TCAGAGGTGG 8800
 CGAAACCCGA CAGGACTATA AAGATACCA GCGTTTCCCC CTGGAAGCTC 8850
 CCTCGTGCAGC TCTCCTGTT CGACCTGCG GCTTACCGGA TACCTGTCCG 8900
 CTTTCTCCC TTGGGAAGC GTGGGCTTT CTCATAGCTC ACGCTGTAGG 8950
 25 TATCTCAGTT CGGTGTAGGT CGTCTGCTCC AAGCTGGGCT GTGTGCACGA 9000
 ACCCCCCGTT CAGCCCGACC GCTGCGCCT ATCCGGTAAC TATCGTCTTG 9050
 AGTCCAACCC GGTAAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT 9100
 AACAGGATTAA GCAGAGCGAG GTATGTAGGC GTTGCTACAG AGTTCTTGA 9150
 GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTG GGTATCTGCG 9200
 30 CTCTGCTGAA GCCAGTTACC TTGGAAGAAA GAGTTGGTAG CTCTTGATCC 9250
 GGCAAACAAA CCACCGCTGG TAGGGTGGT TTTTTGTTT GCAAGCAGCA 9300
 GATTACGCGC AGAAAAAAAG GATCTAAGA AGATCCTTTG ATCTTTCTA 9350
 CGGGGCTCTGA CGCTCAGTGG AACGAAACT CACGTTAAGG GATTTGGTC 9400
 ATGAGATTAT CAAAAGGAT CTTCACTTAG ATCCCTTTAA ATTTAAAATG 9450
 35 AAGTTTTAA TCAATCTAA GTATATATG GTAAACATTGG TCTGACAGTT 9500
 ACCATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCGT 9550
 TCATCCATAG TTGCTCTGACT CCCCCCTGCTG TAGATAACTA CGATACGGGA 9600
 GGGCTTACCA TCTGGCCCA GTGCTGCAAT GATACCGCGA GACCCACGCT 9650
 CACCGGCTCC AGATTTATCA GCAATAAAACC AGCCAGCCGG AAGGGCCGAG 9700
 40 CGCAGAAGTG GTCCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAAATTG 9750
 TTGCGGGGAA GCTAGAGTAA GTAGTGTGCC AGTTAATAGT TTGCGCAACG 9800
 TTGTTGCCAT TGCTACAGGC ATCGTGTGTG CACGCTCGTC GTTTGGTATG 9850
 GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCAGGTTA CATGATCCCC 9900
 CATGTTGTGC AAAAAGCGG TTAGCTCCTT CGGTCTCCCG ATCGTTGTCA 9950
 45 GAAGTAAGTT GGCGCAGTG TTATCACTCA TGTTTATGGC AGCACTGCAT 10000
 AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTCTG TGACTGGTGA 10050
 GTACTCAACC AAGTCATCTC GAGAATAGTG TATGCGGCCA CCGAGTTGCT 10100
 CTTGCCCGGC GTCAATACCG GATAATACCG CGCCACATAG CAGAACTTTA 10150
 AAAGTGTCA TCATTGGAAA ACGTTCTTCG GGGCAAAAC TCTCAAGGAT 10200
 50 CTTACCGCTG TTGAGATCCA GTTCACTGTA ACCCACTCGT GCACCCAAC 10250
 GATCTTCAGC ATCTTTACT TTCAACAGCG TTCTGGGTG ACCAAAAACA 10300
 GGAAGGCAAATGCGCAA AAAGGAAATA AGGGCGACAC GGAAATGTG 10350
 AATACTCATA CTCTTCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT 10400
 ATTGTCAT GAGCGGATAC ATATTGAAT GTATTAGAA AAATAAACAA 10450
 55 ATAGGGGTTTC CGCGCACATT TCCCCGAAAA GTGCCAC 10487

SEQ ID NO:33 (conalbumin polyA)

60 tctgccattg ctgcttcctc tgcccttcct cgtcaactctg aatgtggctt cttcgtact
 gccacagcaa gaaataaaat ctcaacatct aaatgggtt cctgaggtt ttcaagagtc
 gttaaagcaca ttccctcccc agcaccacccctt gctgcaggcc agtggccaggc accaacttgg

ctactgctgc ccatgagaga aatccagttc aatattttcc aaagcaaaat ggattacata
tgcccttagat cctgattaaac aggcgttgt attatctagt gcttcgctt caccaggatt
atcccatattgc ctcccc

5

SEQ ID NO:34 (exemplary antibody light chain sequence)

```
 1 gagtcgtga tgacccagac tccatccccc ctgtctgcct ctctgggaga cagagtcacc
 61 atcagttgca gggcaaatca ggacattagc aattatttaa actggtatca gcagaaacca
121 gatggaactg taaaactcct gatctactac acatcaagat tacactcagg ggtcccata
181 aggttcagtgcagttggc tggaaacagat tattctctca ccattagcaa cctggagcaa
241 caagattttgcacttactt ttgccaacag ggtaatacgc ttccgtggac gttcggtgaa
301 ggcaccaacc tggaaatcaa acgggctgat gctgcaccaa ctgtatccat cttccacacca
361 tccagtgagc agttaacatc tggaggtgcc tcagtcgtgt gcttcttgc aaacttctac
421 cccaaagaca tcaatgtcaa gtggaaagatt gatggcagtgc aacgacaaaaa tggcgtcctg
481 aacagttgga ctgatcagga cagcaaagac agcacctaca gcatgagcag caccctcagc
541 ttgaccaagg acgagttatga acgacataac agctatacct gtgaggccac tcacaagaca
601 tcaacttcac ccattgtcaa qaqcttcaac aqqaatqagt qttaa
```

20

SEQ ID NO:35 (exemplary antibody heavy chain sequence)

1 ctgcgactcgac gacctggcct ggtggcccc tcacagaacc tgccatcac ttgcactgc
61 tctgggtttt cattaaccag ctatggtga cactgggtc gccagcctcc aggaaagggt
121 ctggaatggc tggagtaat atggactggt agaagcacaa ctataattc ggctctcatg
181 tccagactga gcatcagcaa agacaactcc aagagccaag ttttcttaaa aatgaacagt
241 ctgcaaactg atgacacagc catttactac tggcagag ggggtctgat tacgtcctt
301 gctatggact actggggtca aggaacctca gtcaccgtct cctcagccaa aacgacaccc
361 ccatctgtct atccactggc ccctggatct gtcgcctaaa ctaactccat ggtgaccctg
421 ggatgcctgg tcaagggtca tttccctgag ccagtgcacag tgacctggaa ctctggatcc
481 ctgtccagcg gtgtgcacac ctccccagct gtcctgcagt ctgacctcta cactctgagc
541 agctcagtga ctgtccccctc cagcacctgg cccagcgaga ccgtcacctg caacgttggc
601 caccggcca gcagcacaa ggtggacaag aaaattgtgc ccaggattg tactagt

SEQ ID NO: 36 (pTnMCS)

35

1 ctgacgcggcc ctgttagcggc gcattaagcg cgccgggtgt ggtggttacg cgcagcgtga
61 cccytactact tcggcacgcgc cttagccggcc ttcccttcgc ttcttccttcc tccttttcgc
121 ccacgttcgc cggccatcga ttggctattt gccattgtatc acgttgtatc catatcataa
181 ttgttacattt tattatggct catgttccaaac attaccggca tggtagacat tattattggac
241 tagttattaa tagtaatcaa ttacggggtc attagttcat agccccatata tggagtccg
301 cgttacataa cttacggtaa atggccgcgc tggctgaccg cccaaacgacc cccgcccatt
361 gacgctcaata atgacgtatg ttcccatatg aacgcctaata gggactttcc atttgactca
421 atgggtggag tattttacggt aaactggcca cttggcagta catcaagtgt atcatatgcc
481 aatgtacggcc ccttatttgcgc tcaatgacgg taaatggccc gcctggccat atgcccggta
541 catgacccta tgggactttc tcaacttggca gtatcatctac gtatttagtca tcgttattac
601 catgtgtatg cgggttttggc agtacatcaa tgggctgttgc tagcgggtttg actcaacgggg
661 atttccaatg ctccacccca ttgacgtcaa tggggacttttgc tttttggcacc aaaatcaacg
721 ggacttccaa aatgtcgtaa acaacttgcgc cccatgtacg caaatggcg gttagggctgt
781 acgggtggag gtctatataa gcagagctcg ttttagtgcac cgctcagatcg cctggagacg

841 ccatccacgc tgttttgacc tccatagaag acaccggac cgatccagcc tccggggccg
901 ggaacggtgc attggAACgc ggattccccc tgccaagagt gacgtaagta ccgcctata
961 actctataagg cacaccctt tggtcttataat gcatgtata ctgtttttgg ctggggccct
1021 atacaccccccc gtttccctat gctatagggtt atggtagatgc tttagctata ggtgtgggg
1081 attgaccattt attgaccactt cccttattgg tgagatata ttccattact aatccata
1141 atggctttt gccacaacta tcttattgg ctatagccca atactctgtc cttcagagac
1201 tgacacggac tctgttattt tacaggatgg ggtcccttattt attatttaca aattccata
1261 :acaacaacg ccgtcccccg tggcccgagt ttttttttttt catagcgtgg gatctccacg
1321 cgaatctcg gtagctggc cggacatggg ctcttctccg ttagccgggg agcttccaca
1381 :cgacccgtt ggtccctatgc ctccacggc tcatggctc tggccgttcc tttgtccctt
1441 acagtggagg ccagacttag gcacagcaca atggcccca ccaccagtgt gcccacaaag
1501 :ccgtggcg tagggatgt gtctgaaaat gaggctggag atgggctcg cacggctgac
1561 gcagatggaa gacttaaggc aegggcagaa gaagatgcag gcaagctgatg tttgtatc
1621 tgataagagt cagagttac tccctgtgg gtagctgtttt cgggtgggggg cagtgtatc
1681 tgagcagtt tctgttgc cggccggcc accagacata atagtcgaca gactaaacga
1741 ctgtttttt ccattgggtt ttttgcgtt caccgtggg ccatgtgcga atctcgatatt
1801 ttacacgact ctcttacca atttgcggcc gaattacact taaaacgact caacagctt
1861 acgttggctt gccacgcatt acttgactgt aaaacttca ctcttaccga acttggccgt
1921 aatccgcacca cccaaacgcg aaaaaaacat aacatcaaaac gaaatcgaccg atttgtatc
1981 aatgtcataa cccaaacaaaga ggcgtcgat gtataccgtt ggcgtcgat cttttatgtt
2041 tccggcaata cgatggccat tttttttttt gacttgcgtt atattcgta gcaaaaaacga
2101 cttaggtat tgcgagctt cgtcgacta caccgtggg ccatgtgcga atctcgatatt
2161 gctgtcccg tttcagagca atgttcaaaag aaggtctatg accaatttttctt agccgaccc
2221 gcgacatcc taccgttac caccacccgc ctatgtca tttttttttt gtagtgcgtt cttttaaagt
2281 ccattgtt aatccgttga gaagttttttt tttttttttt tttttttttt tttttttttt
2341 gtacaaatgtt cagacttagg aegggaaaac tgaaaaacca tccaaactt atctcgatatt
2401 :catctatgtc actcaaaagac ttttaggttat aagaggctga taaaacgaa tccaaatctca
2461 tgccaaatcc tattgtataa atctcgctt aaggccgaa aaaaatcagcg ctcgacaccc
2521 acttattgtc accaccggc acctaaaatc tactcgatgtt cggccaaaggaa gccatgggg
2581 cttagcaactt acttacatgt tttttttttt tttttttttt tttttttttt tttttttttt
2641 aagcgaatgc agatggaaaac aacccttccg gacttggaaa gttctgcctt cggataggc
2701 ctacccatata gccgaacggc cagtttgcggg cttttttttt tttttttttt tttttttttt
2761 atgcttcaac taacatgttg gcttgcgggc gttcatgtctt tttttttttt tttttttttt
2821 cacttccagg ctaacacatc cagaatcgaa aacgtacttcc tttttttttt tttttttttt
2881 gaatgttgc ggcattttttt tttttttttt tttttttttt tttttttttt tttttttttt
2941 ctactatgtc aaaaattttt tttttttttt tttttttttt tttttttttt tttttttttt
3001 tcttagagcga tccggggatct cggggaaaagc gttttttttt tttttttttt tttttttttt
3061 ctttttttttt aaaaacaaat tactcgttgc tttttttttt tttttttttt tttttttttt
3121 tgcctacatc acaacaaaaaa ctgtttttaac aatgtttttttt tttttttttt tttttttttt
3181 tgaacatattt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
3241 gatgcctatc atgggttggaa atgaacttgc aaaaattttt tttttttttt tttttttttt
3301 agttttttttt aaaaacaaat tttttttttt tttttttttt tttttttttt tttttttttt
3361 aatgtttaattt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
3421 attaagttaa ggttggataca catcttgcata tttttttttt tttttttttt tttttttttt
3481 catttggcactt cccggatgtt acactttttt ctccggctc tttttttttt tttttttttt
3541 agccggataac aatttcacac agggaaaacgc tatgtttttt tttttttttt tttttttttt
3601 aacccttactt aaaaacaaat tttttttttt tttttttttt tttttttttt tttttttttt
3661 agtggatccc cccgggttgc ggaattcgat tttttttttt tttttttttt tttttttttt
3721 gggggggcccg gtacccattt cccggatgtt tttttttttt tttttttttt tttttttttt
3781 :gtgtgtttttt cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
3841 :catccccctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
3901 :acagtgcgcg agcgtttttt gcaatggaa atttttttttt tttttttttt tttttttttt
3961 gctttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
4021 cttttttttt aaaaacaaat tttttttttt tttttttttt tttttttttt tttttttttt
4081 agtccactat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
4141 :atgtggccac tactccggg tttttttttt tttttttttt tttttttttt tttttttttt
4201 :tacaaaaattt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
4261 agggaaagttt atgtatgtt tttttttttt tttttttttt tttttttttt tttttttttt
4321 :aatacatgc gaaaaacccat tttttttttt tttttttttt tttttttttt tttttttttt
4381 :accggcgctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
4441 aatctttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
4501 :gataaaatccat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
4561 ttgaggggtt aacatgttgg tttttttttt tttttttttt tttttttttt tttttttttt
4621 aataatccaa atccagccat tttttttttt tttttttttt tttttttttt tttttttttt
4681 taatggggcaataaaccggg tttttttttt tttttttttt tttttttttt tttttttttt
4741 :ttgtctgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
4801 acccaggcc aaaaaattttt aacacggggaa aacttacccaa cttttttttt tttttttttt
4861 aaaggcaat gcaactactat tttttttttt tttttttttt tttttttttt tttttttttt

SEQ ID NO:37 (chicken ovalbumin enhancer)

45	ccgggctgca	aaaaaatgcc	aggtggacta	tgaactcaca	tccaaaggag	
	cttgcctga	tacctgattt	tcttcaaaact	ggggaaacaa	cacaatccca	caaaacagct
	cagagagaaa	ccatcaactga	tggctacacg	accaaggat	gcaatggcaa	tccattcgcac
	attcatctgt	gacctgagca	aaatgattta	tcttccatg	aatggttgt	tcttcctc
	atgaaaggc	aatttccaca	ctcacaatat	gcaacaaaga	caaacagaga	acaattaatg
	tgtctttcc	taatgtcaaa	attgttagtgg	caaagaggag	aacaaaatct	caagttctgat
50	gtagggttta	gtgattggat	aagaggctt	gacctgttag	ctcacctgga	cttcataatcc
	ttttggataaa	aaaggcttt	tataactttc	aggctccgg	gtctttattc	atgagactgt
	tgggttaggg	acagaccaccc	aatgaaatgc	ctggcatagg	aaaggggcagc	agagccttag
	ctgacccccc	cttgggacaa	gcattgtcaa	acaatgtgt	acaaaactat	ttgtactgt
	ttgcacagct	gtgctgggca	gggcaatcca	ttgccaccta	tcccaggtaa	ccttccaact
	gcaagaagat	tgttgcattac	tctctctgata			

SEQ ID NO:38 (5' untranslated region)

GTGGATCAACATACAGCTAGAAAGCTGTATTGCCTTAGCACTCAAGCTAAAAGACAACTCAGAGTC
ACC

60 SEQ ID NO:39 (putative cap site)
ACATACTAGCTAG AAAGCTGTAT TGCCCTTAGC ACTCAAGCTC AAAAGACAAC TCAGAGTTCA

SEQ ID NO:40 (fragment of ovalbumin promoter - chicken)

GAGGTCAAGT GGTTCTTA CTGTTGTCA ATTCTATTAT TTCAATACA
AACAAAGCT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG

TCCCTCGAAC CATGAAACACT CCTCCAGCTG AATTCACAA TTCCCTCTGTC
 ATCTGCCAGG CCATTAAGTT ATTCACTGGAA GATCTTGAG GAACACTGCA
 AGTTCATATC ATAAACACAT TTGAAATTGA GTATTGTTTT GCATTGTATG
 GAGCTATGTT TTGCTGTATC CTCAGAAAAA AAGTTTGTAA TAAAGCATTC
 5 ACACCCATAA AAAGATAGAT TAAATATTTC CAGCTATAGG AAAGAAAGTG
 CGCTCTGCTC TCACTCTAGT CTCAGTGGC TCCTTCACAT GCATGCTTCT
 TTATTTCTCC TATTTGTCA AGAAAATAAT AGGTACAGTC TTGTTCTCAC
 TTATGTCTG CTCAGCATGG CTCAGATGCA CGTTGTAGAT ACAAGAAGGA
 TCAAATGAAA CAGACTCTG GTCTGTAACT ACAACCATAAG TAATAAGCAC
 10 ACTAACTAAT AATTGCTAAT TATGTTTCC ATCTCTAAGG TTCCCACATT
 TTCTGTTT CTTAAAGATC CCATTATCTG GTGTAACCTG AAGCTCAATG
 GAACATGAGC AATATTTCCC AGTCTCTCT CCCATCCAAC AGTCTGTATG
 GATTAGCAGA ACAGGCAGAA AACACATTGT TACCCAGAAT TAAAAACTAA
 TATTTGCTCT CCATTCAATC CAAAATGGAC CTATTGAAAC TAAAATCTAA
 15 CCCAATCCA TTAAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG
 GGAACCTGTG GGTGGGTAC AATTCAAGGCT ATATATTCCC CAGGGCTCAG
 C

 SEQ ID NO: 41 pTnMCS (CMV-CHOVg-ent-ProInsulin-synPA)
 20 1 ctgacgcgc ctgttagcggc gcattaagcg cggcggtgt ggtgggtacg cgacgcgtga
 61 ccgctacact tgccagcgc cttagccccg ctcccttcgc ttcttcctt tccttcgc
 121 ccacgttcgc eggcatcaga ttggattgtt gccattgtat acgttgcatacataataa
 181 tatgtacatt tatattgtt catgtccaaat attacggca ttgttgcatt gattattgc
 241 tagttatcaa tagtaatcaa ttacgggttc atttagttcat agccatataa tggttgc
 301 cgttacatcaa cttacggtaa atggccgcgc tggtgtacccg cccaaacgcacc cccgccccatt
 361 gacgtcaata atgacgtatg ttcccatatgt aacgccaata gggactttcc attgacgtca
 421 atgggtggag tatttacggt aaactggccca ttggcagta catcaagtgt atcatatgc
 481 aagtacgccc cttattgtacg tcaatgtacgg taaatggccc ggttgcatt atggccatgt
 541 catgacccat tgggactttc ctacttgcgc gtacatctac gtatattgtca tcgttattac
 601 catggtgatg cggttttggc agtacatcaa tggcgtggaa taggggtttg actcacgggg
 661 atttcaagt ctccacccca ttgacgtca tggaggttttgg ttttgcacc aaaatcaacg
 721 ggactttcca aaatgtcgtaa acaatccgc cccattgtacgg caaatggcgtt gtaggcgtgt
 781 a:gtggggag gtctatataa gtagactgtc tttagtgcac cgttagatcg cctggagacg
 841 cgtatccacgc tggtttggc tccatagaag acacggggcgc gatccacggc cccggccgc
 901 ggaacgggtc atttggaaacgc ggattccccc tgccaaaggt gacgtaaatg cccctatag
 961 atctataggg cacacccctt tggctttat gatgtctata ctgttttgg ctggggcc
 1021 atacacccccc gtttccctat gatctatagggtt atgtatagc ttacgtctata ggtgtgggg
 1081 attgaccattt attgaccactt cccttatttttgg tgacgtactt ttccattact aatccatataac
 1141 atgtcttgc tccacaacta tctcttatttttgg ctatatgcctt atactctgtc ttccagagac
 1201 tgacacggc tctgtatattt tacaggatgg ggtcccttattt attatttaca aattccatata
 1261 tacaacaacg cggccccccg tgcccgcaatg ttttattttaa catagcgtgg gatctccac
 1321 cgaatctccg gtacgtgttc cggacatggg ctcttctccg gtacggccgg agcttccac
 1381 tccggccctt ggtcccatgc ctcacggccgc tcatggccg tccggccatgc cttgtccctt
 1441 acagtggagg ccacacttag gcacacgcata atggcccaacca ccacccatgt gccgcacaaag
 1501 gccgtggcc taggttatgt gtcggaaaat gacgtgtggat atgggctcg cacggctgac
 1561 gcatgtggaa gacttaaggc agccgcacaa gaagatgtacg gacgtgtatc
 1621 tgataaaggat cagaggtaac tcccttgcgt gtcgtttaa cgttgggggg cagtgtatc
 1681 tgacgtacat tgggtgtcc cggccgcgc accagacata atagctgaca gactaaacaga
 1741 ctgttccctt catgggtt ttttgcgtt caccgtcgaa ccatgtgcga actctgatatt
 1801 ttacacgact ctcttttacca atttgcggcc gaaatttactat taaaacgactt caacagctt
 1861 acgttggctt gccacgcattt acttgcgtt aaaaactctca ctcttaccgtt acttggccgt
 1921 aacctgccaa cccaaacgcacaa aacaaaacat aacatcaaaac gatgtacccg attgtatgt
 1981 aatctgttccatcccaaaaga ggcactgtt gatatacggtt ggcacgtgtc ctttattttgt
 2041 tccggcaataatccatgttgcgttccatgttgcgtt gatgttgcgtt atattctgtg gaaaaaaac
 2101 ctatgtgttgcgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2161 cgttcccccgc tttcagagca atgttcaaaag aaagctcatg accaatttctt agccgcaccc
 2221 cgtcggccatccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2281 cgtatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2341 gtagacatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2401 tccatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2461 tgccaaatccatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2521 ctttattttgttgcgttccatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2581 ctatgtgttgcgttccatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2641 aacggcaatgc agatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2701 ctatgtgttgcgttccatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt
 2761 atgttgcgttccatgttgcgttccatgttgcgttccatgttgcgttccatgttgcgtt ctttattttgtt atttatgt

5	2821	cacttccagg	ctaacacagt	.cagaatacg	aacgtactct	caacagttcg	cttaggcatt
	2881	gaagttttgc	ggcattctgg	ctacacaata	acaagggaa	acttactcgt	ggctgcaccc
	2941	ctactagtc	aaaatttt	cacacatgtt	tacgccttgg	ggaaattatg	aggggatcgc
	3001	tcttagagca	tccggatct	cggaaaagc	tacgggtacc	aaagggtcct	tttatcatca
	3061	ctttaaaaat	aaaaaaata	tactcgtgc	ctgttatacg	cagaattaa	ttatgatgg
	3121	tgcctacatc	acaacaaaaa	ctgatTTAAC	aaatggttgg	tgccttag	aaagtatatt
	3181	tgaacattat	cttgattata	ttattgataa	taataaaaaac	tttaccccta	tccaagaagt
10	3241	gatgcctatc	attgggtgg	atgaaacttga	aaaaaaattag	ccttgaatac	attactggta
	3301	aggtaaacgc	catttgcage	aaattgtatcc	aagagaacaca	acttaaagct	ttctgcacgg
	3361	aacttataat	ctcggtgacc	ctggactgt	atgaatcccc	taatgtttt	ggtaaaaatc
	3421	attaaggtaa	ggtgataca	cattttgtca	tatgatcccc	taatgttag	ttatgtcaat
	3481	cattaggcac	cccaggctt	acactttatg	cttccggctc	gtatgttg	tggaatttg
15	3541	agcggataac	aattttcacac	aggaaaacagc	tatgaccatg	attacgc	gcgcgcacat
	3601	aacccttact	aaagggaaaca	aaagctggag	ttccacccgc	gtggccggcc	ctctagaact
	3661	agtggatccc	cggggatca	gattggctat	tgggcattcg	atacgttga	tccatcatat
	3721	aataatgtaca	tttatattgg	ctcatgtca	acattaccgc	catgttgaca	ttgattattg
	3781	actagttatt	aatagaatc	aattacgggg	tcattagttc	atagcccata	tatggagttc
20	3841	cgcgttacat	aacttacgtt	aaatggcccg	cctggctgac	cgcacaa	ccccccgcac
	3901	tttgcgttca	taatgcgtt	tgttccca	gtaaacgc	tagggactt	ccattgtacgt
	3961	caatgggtgg	agtattttac	gttaactgc	cattttgcag	catatcaatg	gtatcatatg
	4021	ccaagtaacgc	ccccatattga	cgtcaatgac	ggtaatggc	ccgccttggca	ttatgcacccag
	4081	tacatgac	tatgggactt	tcctacttgg	cagtacatct	acgtattatg	catcgctatt
25	4141	accatgggttga	tgccgttttg	gcagttacatc	aatgggcgtg	gatagcgtt	tgactcacgg
	4201	ggatttccaa	gtctccaccc	cattgtacgtc	aatggggat	tgttttggca	ccaaaatccaa
	4261	cgggacttcc	aaaaatgtcg	taacaactcc	gccccattga	cgcaaaatgg	cggttaggc
	4321	gtacgggtgg	aggcttatata	aaggcagact	cgtttagtga	acgcgtatg	cgccgtggaa
	4381	cgccatccac	gctgtttga	cctccataga	agacacccgg	accgatccag	cctccgcggc
30	4441	cgggaaacgg	gcatttggaa	gcccattttcc	cgtgccaaga	gtgacgtaa	taccgcctat
	4501	agactctata	ggcacaccccc	tttgcgttctt	atgcatgtca	tactgtttt	gggttggggc
	4561	ctataccaccc	cgcgttccctt	atgtatagg	tgatgtttt	gtttagctta	taggtgtgg
	4621	tttattgtacca	tttattgtacca	cttcccttatt	ggtgacgata	ctttccattta	ctataccata
	4681	acatggctct	tttgcacaaac	tatctctatt	ggctatatgc	caataactctg	tccttcagag
35	4741	actgacacgg	actctgtatt	tttacaggat	ggggtcccat	tttatttattt	caaattcaca
	4801	tatatacaca	cgccgtcccc	cgtgtccgc	gttttttta	aaatagcgt	gggatctccca
	4861	cgcgaatctc	gggtgttgtt	tccggacatg	gggttttttc	cgttagcggc	ggaggttccca
	4921	catcgacgg	ctgggtccat	gcctccagcg	gctcatgtc	gttcggcgc	tccttcgtcc
	4981	taacagtgg	ggcccgactt	aggcacagca	caatgc	caccaccatg	gtgcgcacaa
40	5041	aggccgtggc	ggtagggat	gtgtctgaaa	atgagcgttgg	agattgggt	cgcacggctg
	5101	acgcagatgg	aagacttaa	gcagcggcag	aagaatgtc	aggcgtgt	gttgtgtat
	5161	tctgtataga	gtcaggatgtt	actccgttgc	cgtgtgttttgc	aaagggtgg	ggcgtgttag
	5221	tctgtatgtt	actcgttgc	ggcggcgcgc	ccacccagaca	taatagctga	cagactaaaca
	5281	gactgttctt	tttccatgggt	cttttctgc	gtcaccgtcg	ggatccatgg	gttcatctgg
45	5341	cgcagcaacg	atggaaatttt	gttttgcgtt	attcaaggag	ctcaaagtcc	accatgcacaa
	5401	tgagaaatccat	tttgcgttcc	ccatggccat	catgtcgtat	ctagccatgg	tataacttgg
	5461	tgcaaaaacac	agcaggacagg	cacagataat	taagggtttt	cgctttgtata	aacttccacgg
	5521	atccggagac	agtattttgg	ctcgtgttttgc	cacatgtcg	aacgttctact	tttactcttag
	5581	agacatccctc	aaccaaaatca	ccaaacccaa	tgatgtttat	tcgttccatgg	tttgcgttag
50	5641	actttatgtt	gaagagagat	acccaaatctt	ggccgtatcc	ttgcgtgttgc	tgaaggaaact
	5701	gtatagatgg	ggcttggaaac	ctatcaactt	tccaaacgttgc	cgagatcaag	ccagagact
	5761	catcaatcc	ttgggttagaa	gtcggacaaa	ttgttattatc	aaagaaatgtc	tttgcaccaag
	5821	cttcgttggat	tctcaactgt	caatggtttgc	ggtaatggcc	attgttctca	aaggactgttgc
	5881	ggagaaaaata	ttttaaggatg	aaagacacaca	agcaatgtc	ttcagatgt	ctggacaa
55	5941	aagcaaaatct	gtcggatgt	tgttccatgg	ttgtttttat	agatgtggat	caatggcttc
	6001	tgagaaatag	aagatccctgg	agttccatttgc	ttccgttgggg	acaatgtac	tttttgttgc
	6061	tttgcgttgc	gaagtttgc	gttgcgttgc	tttttttttttt	ataatcaat	tttgcgttag
	6121	gactgaatgg	accaggcttca	atgttgcgttgc	agagaggaa	tttttttttttt	tttttttttttt
	6181	catcgatgtt	gaggaaaaat	acaaacttcac	tttgcgttgc	tttttttttttt	tttttttttttt
60	6241	cgtgttttgc	tcttcgttca	atctgttgc	tttttttttttt	tttttttttttt	tttttttttttt
	6301	tcaagtcgtc	catcgacac	atcgacaaaat	caatgtac	tttttttttttt	tttttttttttt
	6361	agcagatgg	ggaggatgttgc	cttcgttgc	tttttttttttt	tttttttttttt	tttttttttttt
	6421	ccttttttgc	atcaagcaca	tcgcaacccaa	tttttttttttt	tttttttttttt	tttttttttttt
	6481	ccggccggat	catcgatgttgc	accaggatgt	tttttttttttt	tttttttttttt	tttttttttttt
	6541	gacgcacccat	catcgatgttgc	accaggatgt	tttttttttttt	tttttttttttt	tttttttttttt
65	6601	tgttttttgc	ttccgttgc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt
	6661	tcacacccat	ttggatgttgc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt
	6721	aagccccc	ggggggcaga	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt
	6781	gggttggatgttgc	ttccgttgc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt
	6841	gaacaatgtt	gttccatgttgc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt

6901 gcctaaaggc cgaattatcg cggccgcctc agaccaggcg cctggatcca gatcaactct
 6961 ggctaataaa agatcagagc tctagagatc tttgtgttgg tttttgtgg atctgctgt
 7021 cctttagtt gccagccatc tttctataaa aatggggaaa ttgcatcgca ttgtctgagt
 7081 ggtgcaactc ccactgtctc ttctataaa aatggggaaa ttgcatcgca ttgtctgagt
 7141 aggttcatc ttattctggg ggggggggtt gggcagcaca gcaagggggga ggattggaa
 7201 gacaatgcg ggcattgtgg ggcattgtgg ggcattgtgg gtcattgtgg gtcattgtgg
 7261 ctctctctct ctctctctct ctctctgtac ctctctcgag gggggcccg gtcattgtgg
 7321 cgcctatag tgagtcgtat tacgcgcgtc cactggccgt cgtttacaa cgtctgtact
 7381 gggaaaaccc tggcgttacc caactataatc gcttgcgc gacatccccc ttgcgcgact
 7441 ggcgtatag cgaagaggcc cgcaccgtc gcccatttccca acatggcgc agcctgaat
 7501 gcaatggaa atttgcgttac ttatattttt gttttttttt gtcattgtgg gtcattgtgg
 7561 cagctcattt tttaaccaat aggccgaaat cggcaaaaatc ctttataat ctttataat
 7621 gaccgagata ggggttggatg ttgttccagt ttggacaag agtccactat taaaacacgt
 7681 ggatccaac gtcacaaaggc gaaaaacccgtt ctatcggtc gatggcccaacttccggga
 7741 tcatgaca agatgtgtat caccatcaac ttatgttggat tttttttttt ctttataat
 7801 ttcatcgtt ctcagggttca accggatata acatccgttc agggaaacgtt atgtatgt
 7861 tttgttggaa aacttactca atggctgggtt atgcataatcg caatacatcg gaaaaacta
 7921 aaagagcttcccgataaaaa agggccatattt attgcttattt accggccgtt ttttttgg
 7981 ttggaaatgaa aataaaatag attagttttt ttggaaatgaa atatccctt atcgtaaaaa
 8041 atgccttccgtt ggggttacaa ggggttcatc atatccgtt gaaataacatc atttgggt
 8101 gaaataacta accgttgc tcttgcgttac tccctcgac ttgggggtt aacatgaagg
 8161 tcatcgatag caggataata atacagtaaa acgctaaacc aataatccaa atccagccat
 8221 cccaaatttgg tagtgaatga ttataaataa cagcaacag taatggccca ataacacccg
 8281 ttgcattgtt aaggcttacc accatccctt gtaaaggacc ttgcattgtt ctctttgtt
 8341 ggatagacat cactccctgt aatgcgcgtt aagcgcgtt accaccaccc aaaaaattt
 8401 aaacaggaa aactaaccatc ctttgcgttca taaacgttca aaaaacccatc gcaactactat
 8461 ctgcataaaa tccgagcgtt actgcgtttt tttcgccat ttgtggctt ttgtggctt
 8521 cacaaggcttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 8581 atgcattatc tcatcgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 8641 ggcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 8701 ttgttccctt tagtggggtaat tttttttttt tttttttttt tttttttttt tttttttttt
 8761 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8821 taaaaggcttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 8881 cgcttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 8941 gagaggccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9001 ggtcgccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9061 agaattcagggttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9121 cccggccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9181 caaaatcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9241 gttttcccttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9301 cctgtccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9361 tctcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9421 gcccggccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9481 cttatcgccatcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9541 tgcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9601 tatctcgcttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9661 caaacaaacc accgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9721 aaaaaggatcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9781 cggccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9841 ctttttttttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9901 tgacgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 9961 atccatagttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10021 tggcccccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10081 aataaaaccc accgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10141 catccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10201 gggccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10261 ttttttttttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10321 aaaaaggatcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10381 atccatcgatcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10441 ctttttttttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10501 gaggttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10561 agtgcgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10621 gagatccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10681 caccagccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10741 ggcgacacccgaaatgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10801 tcagggttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt
 10861 aggggttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgttccgtt

SEQ ID NO:42 (pTnMOD (CMV-CHOVg-ent-ProInsulin-synPA))

1 ctgacgcgcc ctgttagcggc gcattaagcg cggcggtgt ggtggttacg cgcagcgtga
 5 61 ccgtacact tgccagcgc ctagcgcgc ctcccttcgc tttcttcctt tccttcgc
 121 ccacgttcgc cggcatcaga ttggctattt gccattgcat acgttgtatc catatcataa
 181 tatgtacattt tatattggct catgtccaaac attaccgca ttgtgacattt gattttgac
 241 tagtattttaa tagtaatcaa ttacgggttcc attatgttac agccatata tggagtccg
 301 cgttacataa cttacggtaa atggccgcgc tggctgaccc cccaaacgacc cccggccatt
 361 gacgtcaata atgacgtatg ttcccatatg aacgccaata gggacttcc attgacgtca
 421 atgggtggag tattttacgtt aaactgccc cttggcagta catcaagtgt atcatatgcc
 481 aagtacgcgcc cctatttgcac tcaatgacgg taaatggccc gcctggcatt atgcccagta
 541 catgacccctt tggacttttca ttactggca gtacatctca gtattatca tcgttattac
 601 catggatgatg cgggttttgcg agtacatcaa ttggcggttca tggcggttgc actcacgggg
 661 atttccaatgt ctccacccca ttgacgttca ttggagtttgc ttttggcacc aaaatcaacg
 721 ggactttcca aaatgtcgta acaactccgc cccatttgcg caaatggcg gttaggcgtgt
 781 acgttggggat gtctatattaa gcacatgtcg tttatgttgc acgttgcgtt cctggagacg
 841 ccatccacgc ttttttgcac ttccatggaa acacccgggg cgttccggcc tccggcccg
 901 ggaacgggtgc atttggaaacgc ggattccccc tgccaagatg gacgttgcg tccgttccatag
 961 actctatagg cacacccctt tggcttttgcg tttatgttgc ttttttttgc ttttttttgc
 1021 atacacccca gcttccctt gctataggatg atggatatac ttacccctt ggttggggat
 1081 attgaccattt attgaccattt ccccttatttttgc tttatgttgc ttccattactt aatccatataac
 1141 atggcttttttgc cccatccatgtt tttatgttgc ttatgttgc tttatgttgc tttatgttgc
 1201 tgacacggac ttttttgcac ttccatggaa acacccgggg cgttccggcc tccggcccg
 1261 tacaacaacgc ccgtccccccgc ttggcggttgc tttatgttgc tttatgttgc tttatgttgc
 1321 cgaatcttcgg gtacgttgc cggacatggg ctccatgttgc tttatgttgc tttatgttgc
 1381 tccgagccctt ggtccatgttgc ctccatgttgc tttatgttgc tttatgttgc tttatgttgc
 1441 acagtggggat ccacatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 1501 gccgttggcg tagggatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 1561 gcagatggaa gactttaaggc agccggcggaa gaagatgcg gacgttgc tttatgttgc
 1621 tgataagatg cagaggttac ttccatgttgc tttatgttgc tttatgttgc tttatgttgc
 1681 tgacgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 1741 ctgttccctt ccatgggttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 1801 ttacatgattt ctccatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 1861 acgttggctt gccacgcattt acttgcgttgc tttatgttgc tttatgttgc tttatgttgc
 1921 aacatgtccaa cccaaacgcg aacaaaatcatgttgc tttatgttgc tttatgttgc tttatgttgc
 1981 aatccgttccatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2041 tccggcaataatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2101 ctatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2161 gggatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2221 ggcggatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2281 ccatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2341 tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2401 tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2461 tgccaaattt tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2521 acttccatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2581 ctatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2641 aaggttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2701 ctatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2761 atgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2821 cttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2881 gaaatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 2941 ctatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3001 agatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3061 gatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3121 tgacgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3181 attgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3241 aggttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3301 tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3361 tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3421 tgccatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3481 attatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3541 ctatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3601 ccctatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3661 aatccatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3721 aaggttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3781 attgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc
 3841 gtgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc tttatgttgc

SEQ ID NO:43 (pTnMOD(Chicken OVeP+OVg'+ENT+proins+syn polyA))

60 1 ctgacgcgcc ctgttagcggc gcattaagcg cggcgggtgt ggtggtacg cgacgcgtga
61 ccgcctacact tgccagcggc ctagcgcccc ctcctttcgc ttcttcct tcctttctcg
121 ccacgttctcgcc cggcattcaga ttggctatggc cccattgtat acgttgtatc catatccaa
181 tatgtacat tatattggct catgttccaaac attacggcca tggatcattt gatttttgac
65 241 tagtattaa tagtaatcaa ttacgggtc attagttcat agccatata tggagttccg
301 cgttacataa cttacggtaa atggccccggc tggctgaccg cccaaacgacc cccggccatt
361 gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca

5	421	atgggtggag	tat tt acgg	aaactgcctt	cttggcgat	catcaagtgt	atcatatgc
	481	aagtacgccc	c t attgacg	tcaatgacgg	taaatggcc	gc t ggcatt	atgcccagta
	541	catgaccca	tgggacttc	c t acttggc	gtacatctac	gtattatge	tcgcatttac
	601	catggtgat	cgg t tttgc	ag t acatcaa	tgggcgttg	tagcgtttt	actcacgggg
	661	atttccaag	c t ccacccca	tt t gacgtca	ttggagttt	tttggcacc	aaaatcaacg
10	721	ggactttcca	aaatgtcgta	acaactccg	cccattgac	caaatgggg	gttaggcgtgt
	781	acggtgggag	gtctataaa	gcagagctcg	tttagtgaac	cgtcagatcg	cctggagacg
	841	ccatccacgc	tg t tttgacc	tccatagaag	acaccgggc	cgatccagcc	tccgcggccg
	901	ggaacggc	atggaa gc	ggat t ccccc	ttggcaagat	gacgtaagta	ccgcctatag
	961	actttagg	caacaccc	tggcttat	gcatgtca	ctgtttttt	cttggggcc
15	1021	atacaccccc	g t ttccctat	gctatagg	atggatatac	ttagctata	gggtgggtt
	1081	attgaccatt	attgaccat	cc c tatttg	tgacgatact	ttccattact	aatccataac
	1141	atggcttctt	gccacaacta	tcttatttg	ctatgc	atactctgtc	cttcagagac
	1201	tgacacggac	tctgtat	ta c aggatgg	ggtcccaattt	attatttata	aattcacata
	1261	tacaacaacg	cgtccccc	tgcccgc	tttttttt	catagcggtg	gatccacag
20	1321	cgaatctcg	gtacgtgtc	cggacatgg	cttcttcc	gtagcggcg	agttccac
	1381	tccgagcc	ggtcccattc	ctccagcgg	tc t atgg	tcggcagtc	cttgcctt
	1441	acagtggagg	ccagacttag	gcacagcaca	atgcccacca	ccaccagtg	gcccacaag
	1501	gcccgtggcc	taggtatgt	gtctgaaaat	gagcgtgg	attggcgtc	cacggctgac
	1561	gcagatggaa	gacttaagg	agcggcagaa	gaagatcg	gcagctgat	tgtgttattc
	1621	tgat a agat	cgaggat	tcccggtcg	gtgctgtt	cggtgggg	cagtgtatc
	1681	tgagcgtac	tcgttgc	cgcgcgcc	accagacata	atagctgaca	gactaa c aga
	1741	ctgttctt	ccatgggtc	tttctgc	caccgtcg	ccatgtcg	actcgatatt
	1801	tta c acgact	ctt t tttacca	at t tgc	tttttttt	aaaactcgact	caacagctt
25	1861	acgttggct	gcca c gtt	at t tcgact	aaaacttc	cttcttacca	acttggccgt
	1921	aacctgc	ccaaagcg	aacaaacat	aacatcaac	gaatcgcc	attgttagt
	1981	aat t gtcacc	tccacaaa	g c gactcgt	gtatccgtt	ggcatgct	ctttatctgt
	2041	tcgggcaata	cgatgc	tgtactt	gactgg	ttctgt	gcaaaaaacg
30	2101	ctt t atgtt	tg c gagttc	ag t cgacta	cacggctt	ctgttactt	ttat t gagaaa
	2161	gcgttccgc	ttc t agagca	at t tcgaa	aaagctca	accat t ttt	agccgac
	2221	g c gaggatc	taccgat	caccaccc	ctcattgtc	gtgatgct	ctttaaagtg
	2281	ccatggtata	aatccgtt	gaagctgg	ttgtactt	taagtc	aaaggaaaa
	2341	gtacaatatg	cagacctt	agcggaaa	tggaaac	tcagcaactt	acatgatatg
35	2401	tca t atgtc	acttcaagac	tttagctt	aagaggctg	ctaaaacgaa	tccaa t ctca
	2461	tgc c aaat	tattgtt	at t tcgt	aaaggccc	aaatcgac	ctcgacacgg
	2521	acttattgtc	accaccc	ac t taa	tactcag	cggcaaa	gcccatttt
	2581	ctagcaacta	acttac	t g aaattc	acaccc	aactt	tat t tattc
	2641	aagcgaatgc	agattgaag	aac t tc	gactt	gatc	cgactaggc
40	2701	ctacgc	ccgcac	cgactcg	cg	ttttt	gatacttgc
	2761	atgttcc	taacatgtt	gttgcgg	gttcatgc	tcg	aaaaacaaagg
	2821	cacttcc	c t acacat	cagaatc	aacgtact	tcacag	tttagcgt
	2881	gaagtttgc	ggcatttgc	ctacaca	acaagg	acttact	ggctgc
	2941	ctactagtc	aaaattttt	cacatgg	tacgtt	ggaaattt	atgggatc
45	3001	tctagagcg	tccggat	cggggaa	gttgcgt	aaaggt	tttacatca
	3061	ctt t aaaaat	aaaaacaaat	tactcag	ctgtt	taag	ttat t gattt
	3121	tgc c tatc	aaacaaaa	ctgat	aaatgg	tgc	tttttttt
	3181	tgaacattat	ctt t attt	ttat t gata	taat	aaaa	ttat t cc
	3241	gatgc	at t gggtt	at t gaactt	aaaa	attt	ttat t gat
50	3301	agtttacgc	cattgtc	aaat t gtt	aaat	gacca	ttccgcac
	3361	aat t tttatt	ctcg	cgact	aaat	gatcc	ttcttgcac
	3421	attaatgtt	gggtt	at t aatcccc	aaat	ttat	tttttttt
	3481	cattaggc	ccca	tttgc	tttgc	ttat	tttttttt
	3541	agcggataac	aattt	acactt	tttgc	ttat	tttttttt
55	3601	aacccttact	aaagg	aaat t gtt	aaat	gacca	ttccgcac
	3661	atgttac	ccgg	cgact	aaat	gatcc	ttcttgcac
	3721	ctgcac	tac	ttttt	aaat	gggg	tttttttt
	3781	cagagagaa	ccat	tttgc	tttgc	ttat	tttttttt
	3841	attatctgt	gac	tttgc	tttgc	ttat	tttttttt
60	3901	atggaaaggc	aatttcc	tttgc	tttgc	ttat	tttttttt
	3961	tgc	tttcc	ttat	tttgc	ttat	tttttttt
	4021	gtagg	ttt	tttgc	tttgc	ttat	tttttttt
	4081	tttttgc	aaatgt	tttgc	tttgc	ttat	tttttttt
	4141	tgttttaggg	acagaccc	aat t aaat	tttgc	tttgc	tttttttt
65	4201	ctgac	tttcc	tttgc	tttgc	ttat	tttttttt
	4261	ttgcacag	gtgt	tttgc	tttgc	ttat	tttttttt
	4321	gcaagaagat	tgtt	tttgc	tttgc	ttat	tttttttt
	4381	gttgc	tttgc	tttgc	tttgc	ttat	tttttttt
	4441	agacttctt	atgt	tttgc	tttgc	ttat	tttttttt

4501 aaaaatcagt tgataccctg ggttagaca tccagcaaaa aaatattatt tgcaactacca
 4561 tcttgctta agtcctcaga ctggcaagg agaatgtaga tttctacagt atatatgtt
 4621 tcacaaaagg aaggagagaa acaaaagaaa atggactgta ctaaacttca gctagtggta
 4681 taggaagta attctgctta acagagattt cagtatctc tatgtatgtc ctgaagaatt
 4741 atgtgtact tttttccccc attttaaat caaacagtgc tttacagagg tcagaatgt
 4801 tctttactg tttgtcaatt ctatatttc aatacagaac aataacttct ataactgaaa
 4861 tatatttgct attgtatattt atgattgtc ctcgaaccat gaacacttcc caagctgaat
 4921 ttcacaattc ctctgtcattc tgccaggcca ttaagtattt catgaaagat ctttgaggaa
 4981 cactgaagt tcatatcata aacaatattt aaattgagta ttgtttgc ttgtatggag
 5041 ctatgtttt ctgtatccctc agaaaaaaag ttgttataa agcattcaca cccataaaaa
 5101 gatagattta aatattccag ctatagggaa gaaagtgcgt ctgccttca ctctagtctc
 5161 agttggctcc ttccatgc tgccttctt tttgtcaaga aataatagg
 5221 tcacgttctg ttctcaattt tgcctgcct agcatggctc agatgcacgt ttagatataca
 5281 agaaggatca aatgaaacag atctctggtc tgtaactaca accatagtaa taagcacact
 5341 aactaataat tgcattattt gttttccatc tctaattttt ccacatttt ctgttttctt
 5401 aagatccca ttatctgggt gtaactgaag tcaatggaa catgacaat attcccaat
 5461 ctctctccca atccaaacatc cttgtatggat tagcagaaca ggcagaaaac acattgttac
 5521 ccagaattaa aactaatat ttgccttcca ttcaatccaa aatggaccta ttgaaactaa
 5581 aatctaaccctt aatccattt aatgattttt atggctcaaa aggtcaaaact tctgaaggga
 5641 acctgtgggt ggggtcataat tccaggcttccatc tttttccatc ggctcagccca gtggatcaac
 5701 atacagcttag aagctgtatc tgcctttagc actcaagctc aaaagacaac tcaagatcca
 5761 ccatggggctc catcggcgc gcaagcatgc aattttttt tgatgttcc aaggagctca
 5821 aagtccacca tgccaaatggaa aacatcttctt actgccccat tgccatcatg tcaactctag
 5881 ccatggtata cttgggtgc aaagacagca ccaggacaca gataaataag gttgttcgt
 5941 ttgataactt tccaggatttccatc ggagacatca ttgaagctca gtgtggcaca tctgtaaacg
 6001 tttacttccatc aacttagatccatc aatctcaacc aatcaatgtat gtttattctgt
 6061 tcaagcttgc tcaatgtatc tttttttttt tgatgttcc aaggagctca
 6121 aagtgtgtaa ggaactgtat agaggaggct tggaaacctat caacttcaaa acagctgcag
 6181 atcaagccatc agagctcatc aattctctgg tagaaagtca gacaaatggaa attatcgaa
 6241 atgtcttc gccaacttc gtcgttccatc aactctcaacc gtttctgggtt aatgcccattt
 6301 tttccaaagg actgtggggat aatggggat aatggggat aatggggat aatggggat
 6361 gagtgactga gcaagaaac aacactgtgc agatgtatc ccagattttt ttatggag
 6421 tggcatcaat ggcttctgaa aatggggat tccctggatc tccatttgc aatggggat
 6481 tgagcatgtt ggtgtgttgc cctgtatggaa ttcctggatc tgagcatgtt gagatgtt
 6541 tcaacttgc aaaaactgact gaatggacca gtttcaatgtat tttttttttt aatggggat
 6601 aagtgtactt accttcgttccatc aatggggat aaaaatccaa cttccatctt gtttcttgc
 6661 ctatggccatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 6721 agagccatc gatatctcaaa gtttccatc tttttttttt tttttttttt tttttttttt
 6781 gagggatgggtt ggggtcgc ggggttgc ggggttgc ggggttgc ggggttgc
 6841 ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc
 6901 tggcatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 6961 atggccatc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc
 7021 gtatcttgc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc
 7081 accaaccatc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc ggggttgc
 7141 gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7201 agtggggcg gggcccttgc gggcccttgc gggcccttgc gggcccttgc gggcccttgc
 7261 agaaggctgg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7321 actactgc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7381 ggtttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7441 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7501 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7561 catcgatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7621 agggggatgg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7681 cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7741 gggcccttgc gggcccttgc gggcccttgc gggcccttgc gggcccttgc gggcccttgc
 7801 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7861 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7921 gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 7981 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8041 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8101 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8161 gggcccttgc gggcccttgc gggcccttgc gggcccttgc gggcccttgc gggcccttgc
 8221 cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8281 aatggggatgg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8341 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8401 gggcccttgc gggcccttgc gggcccttgc gggcccttgc gggcccttgc gggcccttgc
 8461 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 8521 taacatcattt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt

5	3581	aggggtaac	atgaaggta	tcgatagca	gataataata	cagtaaaacg	ctaaaccaat
	7641	aatccaaatc	cagccatccc	aaattggtag	tgaatgatta	taataaacag	caaacagtaa
	8701	tgggcaata	acacccggtt	cattggtaa	gctcccaat	aatccctgt	aagccacccat
	8761	ctgtatgactc	ttttgttgg	tagacatcac	tcctctgtat	gcaggtaaag	cgatcccacc
	1821	acaggccaaat	aaaattaaaa	caaggaaac	taaccaacat	tcagatataa	acgtctaaaa
	1881	ggcaaatgca	ctactatctg	caataatcc	gaggcgtact	gcccgtttt	cgcccatat
	3941	gtggctattc	ttccctgccc	aaaggcttgg	aatactgagt	gtaaaagacc	aagacccgt
10	9001	atgaaaagcc	aaccatcatg	ctattcatca	tcacgat	tgtatagca	ccacacccgt
	9061	ctggatgtgc	tatcaatg	ctgaaataat	aatcaacaaa	ttgcacgtt	aaataagtga
	9121	tgatataccg	tcagtttt	ttccctttag	tggggtta	ttgcgcgtt	ggcgtatca
	9181	tggtcatagc	tggttccgt	gtgaaattgt	tatccgtca	caattccaca	caacatacg
	9241	gcccggaaagca	taaagtgtaa	agcctgggg	gcctaatgag	tgagctaact	cacattaatt
	9301	gcgttgcgt	cactgcccc	tttccagtc	ggaaacccgt	cgtgcacgt	gcatatgt
	9361	atccggccaa	gcgcgggggg	aggcgggtt	cgtattggc	gcttccgc	ttccctgcgt
	9421	actgactcgc	tgccgttgg	cgltcggtc	cgggcagggc	tatcgtca	ctcaaaaggcg
	9481	gtataatcgg	tatccacaga	atcaggggat	aacgcagggaa	agaacatgt	agcaaaaggc
	9541	cagcaaaagg	ccaggaaccc	taaaaaggcc	gcgttgc	cgtttttca	taggtccgc
	9601	ccccctgacg	agcatcacaa	aaatcgcgc	tcaactcaga	ggtggggaaa	ccgcacagg
	9661	ctataaaagat	accaggcg	tcccccgtt	tgccgttcc	tgccgttcc	tgttccgacc
20	9721	ctggccgtt	cgggatcc	gtccgc	ctcccttgc	gaagcgtgc	gttttctat
	9781	agctcagcgt	gtaggatatc	cagttcggt	tagtgcgtt	gctccaaagct	gggcgtgtgt
	9841	cacgaacccc	ccgttcagcc	cgaccgc	gccttattcc	gtaactatcg	tctttaggt
	9901	aaccccggtaa	gacacgactt	atccgcact	gcgcggcc	ctggtaacag	gattacgaga
	9961	gcggaggat	tagggcggt	tcagagtt	ttgaatgg	ggcctaacta	cggctacact
	10021	agaaggacac	tatgttgg	ctggcgtct	ctgaaaggcc	ttaatcccg	aaaaaaagggt
	10081	ggtagctt	gatccggaa	acaaaccacc	gctgttagc	gtggttttt	tgtttgc
	10141	cagcagatta	cgccagaaaa	aaaaggatct	caagaagatc	ctttgtatct	tttacacgggg
	10201	tctgcacgctc	agtggacgca	aaatcactcg	taaggat	ttgttcatcg	attatcaaaa
	10261	aggatcttca	cctagatcc	ttaaattaa	aaatgaatgt	ttaatcaat	ctaaatgt
30	10321	tatgtataaa	tttgggtga	cagttaaaaa	tgcttaatca	gtgaggcc	tatctcagcg
	10381	atctgtt	ttcgttcatc	catagttgc	tgactccccc	tcgtgttagat	aactacgata
	10441	cgggaggggc	taccatctgg	ccccagtg	gcaatgatac	cgcgagaccc	acgtccaccc
	10501	gctccagatt	tatcagcaat	aaaccaggca	ggccggaaagg	ccggccgc	aagtggctt
	10561	gcaactttat	ccgcctccat	ccagtctatt	aattgttgc	gggaagctag	agtaatgt
	10621	tcggcaggta	atagtttgc	caacgttgc	ggccatgtca	cgccatcg	gggtgtcacgc
	10681	tcgtcggtt	atgtggttc	attcagctt	ggttcccaac	gatcaaggcg	agttacatga
	10741	tcccccatgt	tgtgc	agcggttag	tccttgcgt	ctccgatgt	tgttcaagat
	10801	aagtggccg	cagtgttac	actctatgg	atggcgcac	tgcataatc	tcttactgtc
	10861	agccatccg	taataggttt	ttctgtgact	ggtgtagact	caaccaagtc	atttctggaaa
	10921	tagtgtatgc	ggccggccag	ttgtcttgc	ccggcgtca	tacggataa	tacccgc
	10981	catagcagaa	cattttaaat	gtctatcatt	ggaaaacgtt	cttcggggcg	aaaactctca
	11041	aggatcttac	cgctgttgc	atccagttc	atgtaaacca	ctcggtcc	caactgtatct
	11101	tcagcatctt	ttacttccat	caggttttct	gggtgtggca	aaacacggaa	gcaaaatg
	11161	gcaaaaaagg	gaataaggcc	gacacggaaa	tgttgaat	tcatactt	cctttttca
	11221	tattattgaa	gatattatca	gggttattgt	tcatgtac	gatacatatt	tgaatgtatt
	11281	tagaaaaata	aaacaaatagg	ggttccgc	acattttcc	gaaaagtgc	ac

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/20389

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C07H 21/24; C12N 15/00
US CL : 536/23.2; 800/24

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 536/23.2; 800/24

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 01/14537 A1 (THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY OF AGRICULTURE) 01 March 2001 (01.03.2001), entire document.	1-51
A	SHERMAN, A. et al. Transposition of the drosophila element marnier into the chicken germ line. Nature Biotechnology. November 1998, Vol. 16, pages 1050-1053.	1-51
Y	WO 99/09817 A1 (BIOTECHNOLOGY AND BIOLOGICAL SCIENCES RESEARCH COUNCIL) 04 March 1999 (04.03.1999), page 5, lines 9-10; page 5, lines 4-5; page 7, lines 15-19; page 21, lines 29-31; page 22, lines 10-24.	1,20,11,18-22,28 and 44
Y	US 6,218,185 B1 (SHIRK et al.) 17 April 2001 (17.04.2001), col. 3, lines 14-30; col. 6, lines 33-35; col. 7, lines 61-67; COL. 14, lines 10-18; col. 17, lines 9-11; col. 18, LINES 1-16, col. 22, lines 7-61; claim 9.	1,20,11,18-22,28 and 44
Y	SCHNEIDER, S. An epitope tagged mammalian/prokaryotic expression vector with positive selection of cloned inserts. Gene. 1997, Vol. 197, pages 337-341, specifically, Abstract, lines 6-8, page 339, col. 1, lines 12-15, col. 2, paragraph 1, line 15; Figure 1.	1, 22, 25, 40, 41, 44

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	
"A"	document defining the general state of the art which is not considered to be of particular relevance
"B"	earlier application or patent published on or after the international filing date
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O"	document referring to an oral disclosure, use, exhibition or other means
"P"	document published prior to the international filing date but later than the priority date claimed
"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&"	document member of the same patent family

Date of the actual completion of the international search

01 February 2004 (01.02.2004)

Date of mailing of the international search report

02 APR 2004

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703) 305-3230

Authorized officer

Valerie Bertoglio

Telephone No. (571) 272-0532

INTERNATIONAL SEARCH REPORT

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KOZAK, M. At least six nucleotides preceding the AUG inhibitor codon enhance translation in mammalian cells. J. Mol. Biol. 1987, Vol. 196, pages 947-950, specifically page 947, col. 1, lines 25-33; page 949, col. 2, line 16.	1, 22, 25, 40, 41, 44
Y	WO 97/47739 A1 (BOARD OF TRUSTEES OPERATING MICHIGAN STATE UNIVERSITY) 18 December 1997 (18.12.1997), page 3, lines 29-30, page 6, lines 13-24, page 9, lines 7-34.	10, 11, 18-21, 28, 46